A PROTEOMIC APPROACH TO INVESTIGATE THE RESPONSE OF TEF (*ERAGROSTIS TEF*) TO DROUGHT

By Rizqah Kamies

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Molecular and Cell Biology, Faculty of Science, University of Cape Town, South Africa

August 2015

Supervised by Dr M.S. Rafudeen and Prof J.M. Farrant

Table of Contents

Declaration

I Rizqah Kamies hereby:

(a) grant the University of Cape Town free licence to reproduce the thesis in whole or in part, for the purpose of research;

(b) declare that this thesis titled:

"A proteomic approach to investigate the response of tef (*Eragrostis tef*) to drought"

(i) is my own unaided work, both in concept and execution, and that apart from the normal guidance from my supervisor, I have received no assistance except as stated below:

(ii) neither the substance nor any part of the above thesis has been submitted in the past, or is being, or is to be submitted for a degree at this University or at any other university, except as stated below:

I am now presenting the thesis for examination for the degree of PhD.

Signature: Signed by candidate Signature Removed

Date: 17 August 2015

Acknowledgements

Firstly, I thank my creator, Allah (SWT) by whose decree, my postgraduate studies have been made possible. It is only through the grace and mercy of Allah (SWT) that this doctoral study has reached completion and I acknowledge my dependence on Allah (SWT) alone in all that I do.

The financial assistance of the National Research Foundation (NRF), UCT Postgraduate Funding and the Ernst and Ethel Eriksen Trust towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF and concerned funding bodies. I would like to thank the NRF and UCT Postgraduate Funding, for the financial assistance during my postgraduate studies.

I would like to acknowledge the support and assistance offered to me by my supervisors, Dr M. S. Rafudeen and Prof J. M. Farrant, thank you for your invaluable guidance, supervision and mentorship. I have constantly appreciated your open-door policy, moments to spare and words of encouragement. I would like to thank Dr Gina Cannarozzi and Dr Zerihun Tadele as well as the Tef research group at the Institute of Plants Sciences, University of Bern, Switzerland. Thank you for all your help, support and guidance and for providing access to the available tef resources used in this study. To Gina, thank you for your time, consideration, guidance and unwavering help with analysing the iTRAQ data and understanding the bioinformatics. In addition, I would like to acknowledge Dr Jonas Grossmann and his colleagues at the Functional Genomics Center, Zurich, Switzerland, for the use of their protViz package and his aid in understanding and analysing the iTRAQ data generated in this study. I would like to thank my colleagues in the Plant Stress Lab (PSL), past and present members, for their help, friendship, support and camaraderie throughout my postgraduate studies. Special thanks to Gina, Keren, Hawwa and Humaira for proofreading, editing and being a fresh set of eyes looking over the chapters.

Lastly, to my family and friends, thank you for your constant support, motivation and understanding. To my parents in particular, Nazeem and Fozia Kamies, you have been foundation and centre and have moulded me into the person I am today, I am eternally grateful for that and love you always. To my husband Musa, your constant love, support, patience and understanding has been instrumental in my achievements and helped me reach my goals. I love you.

List of Figures

List of Tables

List of Supplementary Tables and Figures

List of Abbreviations

2-DE	Two-dimensional electrophoresis
2D-DIGE	Two-dimensional difference gel electrophoresis
ACN	Acetonitrile
AWC	Absolute water content
BCA	Bicinchoninic acid
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
DTT	Dithiothreitol
FASP	Filter assisted sample preparation procedure
FBA	Fructose bisphosphate aldolase
FDR	False discovery rate
GLN	Glutamine synthetase
GO	Gene ontology
HPLC	High pressure liquid chromatography
IPG	Immobilised pH gradient
iTRAQ	Isobaric tag for relative and absolute quantitation
KEGG	Kyoto Encyclopaedia for Genes and Genomes
LEA	Late embryogenesis abundant
MDHAR	Monodehydroascorbate reductase
MMTS	Methyl methanethiosulfonate
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MSG	Monosodium glutamate
MU	Monocot unique
PAGE	Polyacrylamide gel electrophoresis
PCD	Programmed cell death

p*I*	Isoelectric points
PMSF	Phenylmethylsulfonyl fluoride
POX	Peroxidase
PTM	Post-translational modification
PVPP	Polyvinylpolypyrrolidone
QTL	Quantitative trait loci
RFLP	Restriction fragment length polymorphism
RIL	Recombinant inbred line
ROS	Reactive oxygen species
RWC	Relative water content
SDS	Sodium dodecyl sulphate
SOD	Superoxide dismutase
TBS	Tris-buffered saline
TCEP	Tris (2-carboxyethyl) phosphine
TE	Tef Extended
TEAB	Triethylammonium bicarbonate
TEM	Transmission electron microscopy
TEU	Tef Extended unique
TFA	Trifluoroacetic acid

Abstract

Eragrostis tef, commonly known as tef, is an important staple food and forage crop indigenous to Ethiopia. Tef plants are highly adaptable to abiotic stress conditions and are able to grow and produce grain yields under a wide range of environmental conditions, particularly under drought stress. In this study, tef plants were subjected to controlled dehydration stress treatment and physiologically characterised using relative water content (RWC), electrolyte leakage and chlorophyll fluorescence measurements, to establish critical water content stages for investigation of changes to the tef proteome in response to dehydration stress. Physiological testing showed tef viability to be retained to 30% RWC, however, further water loss to below 30% RWC, resulted in total loss of viability. Physiological characterisation with dehydration treatment showed a maximum leakage rate of 780 $\mu S.min^{-1}.gdw^{-1}$ and complete photosynthetic disruption with F_v/F_m and ϕ_{PSII} values decreasing to 0.2, below 30% RWC. Additionally, ultra-structural analysis using transmission electron microscopy showed extensive damage to the subcellular organisation of tef plant cells at water contents below 30% RWC. Based on these physiological data, it was decided to investigate the proteome of tef leaf dehydrated tissues at 50% RWC, as a non-lethal dehydration stress, as compared to hydrated tissues at 80% RWC. Proteomic analyses using iTRAQ mass spectrometry coupled to peptide OFFGEL fractionation and appropriate database searching with the Tef Extended and *Liliopsida* databases enabled the generation of three dataset results. These datasets, each contained a substantial amount of database matched proteins, where 5727 proteins for the Tef Extended (TE), 2656 proteins for the Tef Extended unique (TEU) and 4328 proteins for the Monocot unique (MU) datasets, were identified. Statistical analyses on peptide relative quantification values showed differential regulation of 211 proteins for the TE dataset, 111 proteins for the TEU dataset and 174 proteins for the MU dataset, in response to dehydration stress. A reciprocal BLAST search through the use of OrthoMCL with all three differentially regulated datasets (foregrounds) showed the TE foreground to provide the most comprehensive total protein coverage for further bioinformatics inference. Bioinformatics analysis using the programs Mercator, MapMan and Blast2GO showed a diverse range of biological processes, where functional enrichment of GO-terms involved in biotic and abiotic stress response, signalling, transport, cellular homeostasis and pentose metabolic processes were enriched in tef high-abundance proteins. GO-terms linked to ROS producing processes such as photosynthetic reactions, cell wall catabolism, manganese transport and homeostasis, the synthesis of sugars and cell wall modification were enriched in tef low-abundance proteins. Additionally, KEGG analysis was used to observe tef proteins mapped to various biological pathways, of which the stress-responsive pathways, glutathione metabolism and ascorbate and aldarate metabolism were analysed in depth. Furthermore, biological validation of a few high-abundance proteins generated from iTRAQ analysis in the form of western blotting and relevant enzyme assays were conducted. The results showed the proteins fructose-bisphosphate aldolase (FBA), glutamine synthetase (GLN), functioning in plant maintenance

processes as well as the stress-protective antioxidant proteins, monodehydroascorbate reductase (MDHAR), peroxidase (POX) and superoxide dismutase (SOD) to be accumulated and further support iTRAQ findings.

To date, this is the first study that has investigated the proteome profile of tef in response to dehydration.

Chapter 1: Introduction

1.1 Tef crop general information

Eragrostis tef (Zucc.) Trotter is a domestic wild grass belonging to the Poaceae family, sub-family *Chloridoideae* (*Eragrostoideae*), tribe *Eragrostidae*, sub-tribe *Eragrostae*, and genus *Eragrostis* (Costanza *et al.*, 1979; Stallkneecht *et al.*, 1993; Assefa *et al.*, 2011). *E. tef,* more commonly known as tef (here on referred to as tef), is an indigenous African crop, whose grain is used as a source of income to many resource-poor subsistence farmers and as a staple food source for many low-income consumers (Tadele and Assefa, 2012). Tef is one of the most important cereal crops endemic to Ethiopia where it is annually cultivated over 3 million hectares of land, to provide food for over 70% of the 80 million Ethiopian people (CSA, 2013; Cannarozzi *et al.*, 2014). Tef grain is used in a variety of traditional dishes as a flat pancake-like spongy bread called "*injera*" or in slightly fermented or unfermented breads, "*kita*" and "*anebabero*" (Assefa *et al.*, 1999; 2001; 2011). Tef is also cooked and consumed as porridge and has been known to be used in the brewing of the native beer, "*talla*" (Assefa *et al.*, 2011). In addition to being predominantly used for human consumption, tef straw is used as a fodder for livestock and a supplement to building material by acting as a reinforcing agent in mud bricks (Tatham *et al.*, 1996; Woyessa and Assefa, 2011). In terms of sustainability, the whole crop plant is utilised for Ethiopian livelihood (Tatham *et al.*, 1996).

There are approximately 350 known species in the genus *Eragrostis* (Costanza *et al.*, 1979), where tef is the only species cultivated as a cereal crop. Tef and finger millet (*Eleusine coracana* L.) represent the only two species in the sub-family *Chloridoideae* whose grains are utilized for human consumption (Assefa *et al.*, 2011). Tef seeds are small in size (0.2 -0.3 mg per kernel), ranging from an opaque white to brown and red in colour (Ayele, 1999; El-Alfy *et al.*, 2012). The seeds have low fat and highly nutritious mineral contents such as calcium, phosphorous, iron, copper, aluminium, barium and thiamine (Stallkneecht *et al.*, 1993; Ayele, 1999). The nutritive value of tef grain is more advantageous than some of the major cereals, such as maize, wheat, sorghum and barley especially with respect to the zinc, copper, manganese and lysine content, with the latter often being limiting in other cereals (Ayele, 1999; Baye, 2014). The high iron content in particular, has been suggested as an explanation for the absence of iron-deficiency in the Ethiopian population (Mamo and Parsons, 1987) and is more prevalent in the darker coloured brown and red seeds (El-Alfy *et al.*, 2012).

Tef grain has also become the preferable food source for gluten-sensitive individuals due to its gluten-free index and for those suffering from celiac disease due to the lack of T-cell stimulatory peptides, which affect celiac patients (Spaenij-Dekking *et al.*, 2005). For those individuals suffering from diabetes, the consumption of tef grain is highly recommended, due to being comprised of complex carbohydrates

with slow digesting starch (Baye, 2014). As a forage crop, tef serves as an important cattle feed during the long dry seasons of the year (Assefa *et al.*, 1999). It is the preferable straw for cattle in comparison to straw from other cereals due to its high digestibility (65%) and relatively low protein content (1.9-5.2%) (Tefara and Belay, 2006).

Although tef crop cultivation and diversity is mostly constrained to Ethiopia (Vavilov, 1951), the crop has been cultivated on a very small scale in Eritrea and recently in USA, the Netherlands and Israel (Assefa *et al.*, 2011). In other countries, such as South Africa, India, Kenya, Uganda, Mozambique and Pakistan, tef straw is mostly used as an animal feed or pasture crop (Lester and Bekele, 1981; Tatham *et al.*, 1996; Assefa *et al.*, 1999). At present, the Ethiopian Institute of Biodiversity (EIB) houses approximately 5169 diverse tef germplasm accessions collected from all across the country that serves as a diverse resource pool used by both national and international research groups for tef plant breeding and crop improvement projects (Tesema, 2013; Assefa *et al.*, 2015).

1.2 Tef morphology and plant growth

Tef can be classified as an annual tufted grass of medium height growing from 150 to 200 cm tall in the field (Tefara and Belay, 2006). It has a shallow, fibrous root system and short vegetative life cycle of 2 to 6 months (Jones *et al.*, 1978; Ayele *et al.*, 1996). Tef is an allotetraploid (2n = 4x = 40 chromosomes) with an estimated genome size of 730 Mbp, is self-pollinating in nature and possesses a C4 cycle photosynthetic pathway (Jones *et al.*, 1978; Ayele *et al.*, 1996; Ketema, 1997; Assefa *et al.*, 2003). In addition, tef leaf tissues, being a grass, are long and thin (Fig. 1A), with Kranz type anatomy, where vascular bundles are surrounded by bundle sheath cells in a circular orientation (Takele *et al.*, 2001; Assefa *et al.*, 2011). The stems of tef are long and thin in stature and mostly grow in an ascending direction, however, some cultivars are known to bend or elbow (geniculate) as growth proceeds (Ketema, 1997; Assefa *et al.*, 2011). It has a panicle type of inflorescence (Fig. 1B), with a shape that ranges from open and loose to compact (whip-like) which branch out into spikelets containing both male and female sexual organs for self-fertilisation (Ketema, 1997; Assefa *et al.*, 2011).

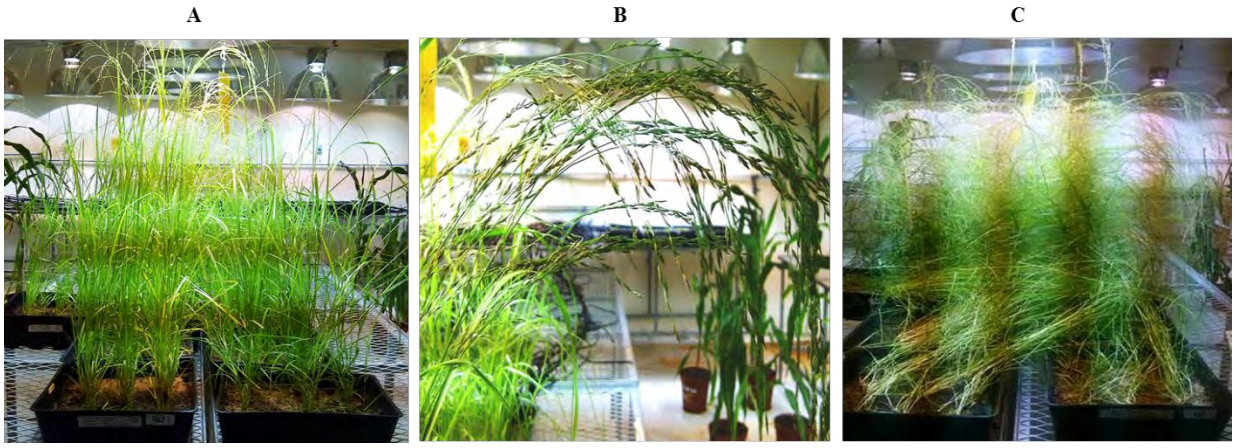

Fig. 1 Matured tef plants (A) showing panicle type inflorescence for seed formation (B) and stems susceptible to lodging (C).

Due to its long and thin stem, tef is often subjected to lodging (Fig. 1C), as a consequence of environmental conditions such as wind, rain and application of nitrogen fertiliser (Tadele *et al.*, 2010; Jöst *et al.*, 2014). Lodging occurs when stems cannot support themselves and are permanently displaced from their upright positions (Tadele *et al.*, 2010) which has been known to cause grain yield losses of more than 30% during harvesting (Girma *et al.*, 2014).

Tef plants are highly adaptable to environmental changes and are able to grow under a wide range of soil types, climatic conditions and at differing altitudes ranging from 1000 to 2500 m above sea level (Costanza *et al.*, 1979; Ayele *et al.*, 1996; Tefara and Belay, 2006). Due to its continuous, extensive cultivation in the subtropical regions of Ethiopia, tef has shown versatile adaptation to both drought and water-logged soils and, over generations, has developed some resilience to limited water conditions (Ayele *et al.*, 1996; Assefa *et al.*, 2003; 2011). During drought conditions, when crops such as maize, wheat and sorghum cannot be maintained under a limited water supply (Tefara and Belay, 2006; Tadele *et al.*, 2010; Cannarozzi *et al.*, 2014), tef is often grown as an insurance or rescue crop, to sustain the Ethiopian population (Shiferaw and Baker, 1996; Ketema, 1997).

1.3 Tolerance to biotic and abiotic stresses

Tef has shown tolerance to a number of biotic and abiotic stresses (Ketema, 1997; Tadele *et al.*, 2010; Cannarozzi *et al.*, 2014). The main biotic stresses affecting tef productivity are caused by species of fungi, of which few have had an economic impact, mostly during specific growth and production years or in certain local regions (Tefara and Belay, 2006). The most growth and grain limiting abiotic stresses are caused by environmental conditions such as drought, water-logging and increased soil acidity and salinity (Tadele *et al.*, 2010; Cannarozzi *et al.*, 2014).

1.3.1 Biotic stresses

In comparison to other cereal crops grown in Ethiopia, tef has shown relative tolerance or resistance to biotic stress conditions caused by the attack of pests, insects and weeds (Stallkneecht *et al.*, 1993; Ketema, 1997; Tefara and Belay, 2006). Previous studies in Ethiopia have identified 22 species of fungi and 3 pathogenic nematodes associated with tef (Bekele, 1985; Stallkneecht *et al.*, 1993). Of the fungi and pests known to causes disease in the humid areas of Ethiopia and affect tef productivity, leaf rust (*Uromyces eragrostidis*), head smudge (*Helminthosporium miyakei*) and damping off (*Drechslera* spp. and *Epicoccum nigrum*), are the most important (Stallkneecht *et al.*, 1993; Tefara and Belay, 2006). The use of fungicides have been shown to be effective in limiting fungal diseases under experimental conditions, however, few have been used in field applications (Tefara and Belay, 2006). The pests or insect species known to attack tef under field conditions include: Wollo-bush cricket (*Decticoides brevipennis*) acting on seeds and seedlings, red tefworm (*Mentaxya ignicollis*), tef epilachna, and tef black beetle (*Erlangerius niger*) affecting inflorescence structures (Stallkneecht *et al.*, 1993; Tefara and Belay, 2006). Although biotic stresses have been known to cause grain losses and are a growing concern to farmers and breeders alike, the loss of tef grain due to abiotic stress factors account for more.

1.3.2 Abiotic stresses

The major abiotic stress factors affecting tef growth and production include drought, soil salinity and acidity (Tadele *et al.*, 2010). While some research has shown that different varieties of tef exhibit relative tolerance to increased salinity (Asfaw and Dano, 2011) and soil acidity (Abate *et al.*, 2013), the majority of studies have reported on tef tolerance to drought stress (Degu *et al.*, 2008; Kreitschitz *et al.*, 2009; Mengistu, 2009; Degu and Fujimura, 2010; Ginbot and Farrant, 2011; Shiferaw *et al.*, 2012b).

1.3.2.1 Salt stress

Tef has been subjected to increased salinity in the lowland and Rift Valley areas in Ethiopia, especially the Awash valley and lower plains (Asfaw and Dano, 2011). Asfaw and Dano (2011) investigated the effects of increased salinity on tef yields and tef components by screening 15 lowland tef genotypes (10 accessions and 5 varieties) at different salinity levels. They found grain yield per main panicle (GY/MP) to be the most affected by increased salinity and, although there were differences in genetic variation among tef varieties and accessions, salt tolerance was observed in accession 237186 and variety DZ-Cr-37 (Tsedey) genotypes (Asfaw and Dano, 2011). Because increased soil salinity conditions affecting grain yield during cultivation are a growing concern in the general areas of Ethiopia, particularly in the Awash valley, the authors have encouraged further investigations to help alleviate the problem (Asfaw and Dano, 2011).

1.3.2.2 Soil acidity

In Ethiopia, under half of the total land area (41%) contains acidic soils of which a third (33%) has been shown to contain high aluminium concentrations (Schlede, 1989; Abate *et al.*, 2013). A recent study involving the observation of different varieties of tef to strongly acidic soil (pH 3.94 and acid saturation

of 78%), was conducted to assess the quantitative tef root length, shoot length, root dry weight and shoot dry weight response (Abate *et al.*, 2013). By using the aluminium tolerant, weeping lovegrass (*Eragrostis curvula* (Schrad.) Nees) variety, Ermelo, as a control, the authors were able to evaluate the response of tef to highly acidic and aluminium-toxic conditions. In general, all tested tef varieties were negatively affected by high acid and aluminium exposure by displaying stunted shoot growth and root pruning in comparison to the control Ermelo variety (Abate *et al.*, 2013). However, among the tested tef varieties (Dima, Emmerson and SA Brown, to name a few), the brown seeded grain variety, Dima, showed a high tolerance for all growth parameters tested under high salt and aluminium soil conditions (Abate *et al.*, 2013).

1.3.2.3 Drought stress

Tolerance to drought stress can be defined as the ability of plants to grow, develop and produce sufficient yields under a limited water supply as a consequence of periodical, environmental or simulated drought conditions (Turner, 1979; Fleury *et al.*, 2010). In most parts of Ethiopia, tef is grown under non-irrigated field conditions during the seasons June to September and February to May (Takele, 1997). As a result, tef crops are regularly subjected to dry-spells where rainfall is limited and yield productivity is affected (Takele, 1997; Mengistu, 2009). Although tef is well suited to growth and development in semi-arid areas often prone to drought conditions (Ketema, 1997; Kreitschitz *et al.*, 2009; Cannarozzi *et al.*, 2014), water-deficit stress or environmental drought is one of the main limiting factors of tef productivity (Degu *et al.*, 2008; Degu and Fujimura, 2010).

Previous studies observing the effect of drought on tef leaves have reported a morphological change in leaf structure (Takele, 1997; Balsamo *et al.*, 2006; Degu *et al.*, 2008), where leaf tissues were reduced in size and area, instead of displaying leaf shedding or death (Shiferaw and Baker, 1996; Takele, 1997). Leaf rolling or the curling inwards of leaves to avoid excessive moisture loss was hypothesised to be an inherent adaptive characteristic of tef to drought conditions (Mengistu, 2009). In a study investigating the physiological responses of tef during different stages of development to drought stress, leaf rolling was accompanied by reduced net CO_2 assimilation rates and reduced photosynthetic and transpiration rates, which differed depending on the developmental stage of tef plants (Mengistu, 2009).

Balsamo *et al.* (2006) investigated leaf tensile properties (leaf behaviour during mechanical stress caused by water loss) of three *Eragrostis* grass species, the drought-sensitive *E. capensis*, the moderately drought tolerant *E. tef* and the drought tolerant *E. curvula*, during dehydration stress. They found a positive correlation of leaf tensile strength with increased dehydration stress (Balsamo *et al.*, 2006). Leaf tensile strength values were the highest for the drought-tolerant *E. curvula*, followed by the moderately drought-tolerant *E. tef* and lastly *E. capensis*, whose leaf tensile strength values were the lowest (Balsamo *et al.*, 2006). The increase in leaf tensile strength was positively correlated with leaf architectural and mechanical changes in cell wall chemistry where ultra-structural studies showed increased lignification of

bundle sheath cells for both *E. tef* and *E. curvula* (Balsamo *et al.*, 2006). The changes in leaf architecture have been proposed to play an adaptive role in stabilisation of the lamina in the *Eragrostis* species, *E. tef* and *E. curvula,* during periods of drought when loss of internal water occurs (Balsamo *et al.*, 2006).

Further physiological investigations by Ginbot and Farrant (2011), where white- and brown-seeded tef varieties were compared to the resurrection grass species, *Eragrostis nindensis*, during a dehydration/rehydration cycle were performed. The study was conducted to observe a better performing crop variety under water-limiting conditions (Ginbot and Farrant, 2011). A decrease in transpiration rates, photosynthetic potential and increased electrolyte leakage rates were observed in tef brown- and white-seeded varieties when relative water contents (RWCs) decreased to 43 and 39% RWC, respectively (Ginbot and Farrant, 2011). Further ultra-structural studies showed damage of membranes and cellular organelles at the water contents, 43 and 39% RWC, for brown- and white-seeded tef varieties, respectively (Ginbot and Farrant, 2011). However, dehydration-induced damage was reversed in the tef brown-seeded variety upon re-watering while the tef white-seeded variety, was unable to recover (Ginbot and Farrant, 2011). Below 30% RWC, however, irreparable damage occurred to the entire subcellular organisation causing loss of plant cell viability (Ginbot and Farrant, 2011). This suggests that the tef brown-seeded variety was able to remain viable until water contents above 30% RWC and is more tolerant to internal water loss in comparison to tef white-seeded variety (Ginbot and Farrant, 2011).

While studies have been conducted on the leaf vegetative tissues of tef in response to dehydration stress (Takele, 1997; Balsamo *et al.*, 2006; Ginbot and Farrant, 2011), only few have focused on the general properties of tef seeds (Bekele *et al.*, 1995; Tatham *et al.*, 1996; Zewdu and Solomon, 2007; Belay *et al.*, 2009; El-Alfy *et al.*, 2012). In a study investigating the structural and physiological adaptations of tef seeds to drought conditions (Kreitschitz *et al.*, 2009), pericarp epidermal cells producing a layer of slime were detected around the fruits of tef and a closely related species, *Eragrostis pilosa* (Kreitschitz *et al.*, 2009). The layer of slime found in the inner cell wall was found to contain pectins that have been proposed to be responsible for water absorption by quickly hydrating and swelling when water is available and thereby maintaining a layer of moisture around the diaspore during seed germination (Kreitschitz *et al.*, 2009). In addition, the slime layer has been proposed to act as an adhesive to dry soil conditions and furthermore, has been suggested to be an adaptive characteristic of *Eragrostis* species under drought conditions or in environments with limited water availability (Kreitschitz *et al.*, 2009).

Among the tef organs investigated for drought tolerance mechanisms, the adjustment of tef root tissues to changing soil moisture contents as a consequence of drought conditions have been reported (Ayele *et al.*, 2001; Degu *et al.*, 2008; Degu and Fujimura, 2010). A deep, penetrating and well-established root system has been suggested to be beneficial for improved drought tolerance, by increasing the ability of plants to mine deeply stored soil water (Ludlow and Muchow, 1990; Wu and Cosgrove, 2000; Degu *et al.*, 2008;

Degu and Fujimura, 2010). The effect of increasing root lengths during water-limiting conditions has been observed in a few cereal crops such as rice, soybean and sorghum (Merrill and Rawlins, 1979; Hoogenboom *et al.*, 1987; Fukai and Cooper, 1995; Degu *et al.*, 2008). Although patterns of root elongation, such as root length under moisture limiting conditions, vary between tef cultivars (Ayele *et al.*, 2001), the primary root tissues of the tef cultivars, Kaye Murri and Ada, were shown to elongate by 34.6 and 35%, respectively, when exposed to drought conditions in comparison to hydrated controls (Degu *et al.*, 2008). The regulation of root elongation has thus been suggested to be an essential adaptive mechanism in response to drought stress in tef (Degu *et al.*, 2008; Degu and Fujimura, 2010).

1.4 Genomics research in tef

Over the last few years, tef has benefited from studies using a diverse array of genetic and genomic tools (Bai *et al.*, 2000; Zhang *et al.*, 2001; Assefa *et al.*, 2003; Ingram and Doyle, 2003; Yu *et al.*, 2006; 2007; Zeid *et al.*, 2011; 2012; Zhu *et al.*, 2012). Although generally considered to be an under-researched or orphan crop in terms of genetic manipulation and improvement (Tadele and Assefa, 2012), studies on tef have provided information on phylogeny, phenotypic and genetic diversity as well as other molecular characteristics (Girma *et al.*, 2014). Most of the studies conducted, however, have been tailored to the generation of molecular tools to assist marker assisted breeding projects for tef growth and improvement under a variety of growth limiting conditions.

1.4.1 Breeding approaches in tef

The breeding of tef for desirable characteristics using scientific approaches was initiated in the 1950s (Assefa *et al.*, 2011). A pioneering study by Ebba (1975), approximately 40 years ago, identified 35 distinct tef ecospecies based on morphological and phenotypical characteristics such as grain and lemma colour, panicle form and ramification, plant and spikelet size and time to maturity (Assefa *et al.*, 1999; Plaza-Wüthrich and Tadele, 2012). Following this study, tef has been enhanced by a broad range of phenotypic diversity characteristics through breeding approaches (Tadesse, 1993; Assefa *et al.*, 1999; 2001; 2003; Tefera *et al.*, 2003). The general breeding targets of tef have been to improve tef germplasm (accession and variety) resources, enhance the scientific knowledgebase of tef crops and to develop tef cultivars for better growth in different climates, landforms and soil conditions (agro-ecological zones) (Assefa *et al.*, 2011). Other specific breeding targets such as increasing tef grain yield and quality, improving lodging resistance and increasing tolerance to drought or water-deficit conditions have also been initiated (Assefa *et al.*, 2011).

Among the specific targeted breeding experiments initiated in tef, the observation of tef drought-tolerance mechanisms was recently advanced with more sophisticated methods using genetic mapping. The generation of genetic maps allows the position of molecular markers and relative quantitative trait

loci (QTL) linked to recombinant frequency on chromosomes to be identified and can subsequently pinpoint genes of interest responsible for specific traits or characteristics (Assefa *et al.*, 2015). A restriction fragment length polymorphism (RFLP) linkage map of tef was generated using 116 recombinant inbred lines (RILs) from a cross with the tef cultivar, Kaye Murri, and close relative *E. pilosa* (Zhang *et al.*, 2001). The RFLP linkage map showed a fair amount of interesting polymorphisms that prompted subsequent gene mapping investigations (Zhang *et al.*, 2001; Chanyalew *et al.*, 2005; Yu *et al.*, 2006; Zeid *et al.*, 2011). The ensuing QTL investigation by Yu *et al*. (2007), identified 99 QTLs for 19 agronomically important traits that could be further used for desirable marker assisted breeding (Girma *et al.*, 2014; Assefa *et al.*, 2015).

Following this cornerstone in tef breeding research, Degu and Fujimura (2010) investigated QTLs responsible for plant height and primary root length in relation to well-watered (hydrated) and water-stressed (drought) environments. They studied a population of 96 RILs derived from a cross between the parental tef cultivars, Kaye Murri and *E. pilosa* based on the results observed by Yu *et al*. (2007). The authors found significant changes in plant height measurements and an increase in primary root length measurements for both parents (Kaye Murri and *E. pilosa*) and RIL plants under water-stressed conditions, where Kaye Murri exhibited larger plant height differences and longer primary root length measurements (Degu and Fujimura, 2010). The increase in primary root growth mechanism during drought conditions (as mentioned earlier) could be an adaptive morphological response of tef to drought whereby roots continue to grow causing repression of plant shoot growth (Wu and Cosgrove, 2000; Degu and Fujimura, 2010). In addition, QTLs for RILs related to plant height and primary root length under both hydrated and drought conditions were identified that could potentially be used in future QTL-aided breeding projects with tef (Degu and Fujimura, 2010).

Although tef has been shown to grow in and adapt to drought conditions, not all tef varieties exhibit drought tolerance and considerable variation exists between varieties (Takele, 1997; Degu *et al.*, 2008). Furthermore, tef varieties respond differently to water-limitations during different stages of growth and development (Mengistu, 2009). A few of the tef varieties investigated that exhibit tolerance to drought and maintain considerable grain yields under water-limiting conditions, include the locally grown varieties: Abat Keyi, Abat Nech, Kobo, Wofey and the breeding improved varieties Tsedey (cultivar DZ-Cr-37), Quncho (DZ-Cr-387) and Dukem (DZ-01-974) (Mengistu and Mekonnen, 2012; Shiferaw *et al.*, 2012a, b). These local varieties, that have been reported to produce better grain yields under water limiting conditions, are due to cultivation in areas often subjected to drought, resulting in their enhanced natural adaptation to water-deficit (Mengistu and Mekonnen, 2012).

1.4.2 Tef genome and transcriptome sequencing

The recent sequencing of the tef genome and transcriptome of the improved variety Tsedey (DZ-Cr-37) is one of the most important achievements towards enhancement of tef growth characteristics. (Cannarozzi *et al.*, 2014; Assefa *et al.*, 2015). The availability of sequence data for tef is an important resource for further 'omic' (genomic, transcriptomic, proteomic and metabolomic) investigations, particularly for highlighting genes that are potentially transcribed and translated into proteins of interest functioning in stress response. The understanding and further knowledge gained from these genes, proteins and metabolites and their respective roles in stress response could facilitate enhancement of tef tolerance to abiotic stress factors particularly tolerance to drought stress.

A few of the more well-known genes linked to drought responses have been detected in the tef genome, these include: *DREB1A, ERD1, SAL1, SNAC1* and *LEA3* (Cannarozzi *et al.*, 2014), that have previously been shown to play a role in drought stress response in plants such as *Arabidopsis*, wheat and rice (Nakashima *et al.*, 1997; Oh *et al.*, 2005; Hu *et al.*, 2006; Xiao *et al.*, 2007; Wilson *et al.*, 2009; Manmathan *et al.*, 2013; Cannarozzi *et al.*, 2014). However, a complete genome set alone is insufficient to elucidate biological function (Pandey and Mann, 2000; Agrawal and Rakwal, 2006; Qureshi *et al.*, 2007). This is because proteins mediate biological processes and are the epigenetic agents involved in almost all biological activities within the cell, particularly by acting as the direct effectors of plant stress response (Pandey and Mann, 2000; Patterson and Aebersold, 2003; Kosová *et al.*, 2011). For further improvement of tef growth and productivity under abiotic stress conditions (such as drought), high-throughput proteomic techniques in combination with the sequenced tef genome and transcriptome could be used for the identification and characterisation of stress responsive proteins.

1.5 Proteomics

Proteomics has become an essential tool for analysing the whole or specific protein complement present in a particular tissue, organ, cell or organelle (Agrawal *et al.*, 2005; Benkeblia, 2011). In recent years, plant proteome analysis has improved due to the evolution of new high throughput techniques resulting in the generation of high quality data with continuous improvements made in sample preparation, protein separation, mass spectrometry and protein search algorithms (Thelen, 2007; Benkeblia, 2011). These improvements have been complemented and strengthened by genome sequencing and annotation (Pandey and Mann, 2000; Agrawal and Rakwal, 2006).

In plant cells, biotic or abiotic stress conditions in most situations induce alterations in gene expression which cause sequential effects in metabolic processes and cellular protein abundance changes within the affected tissues (Kosová *et al.*, 2011; Nanjo *et al.*, 2011). The measurement of these protein abundance

changes and post translational modifications allow key proteins and biological processes to be highlighted for further investigation (Baginsky, 2009; Kosová *et al.*, 2011; Vanderschuren *et al.*, 2013). In addition, the examination of how protein profiles change in response to stress conditions is critical for the understanding of drought tolerant phenotypes and molecular mechanisms involved in stress tolerance or adaptation (Nanjo *et al.*, 2011).

The analysis of proteomic changes in response to abiotic stress factors such as drought is advantageous over transcript-based techniques especially for large-scale study of associated molecular changes (Benešová *et al.*, 2012). This is potentially due to the high conservation of protein sequences in comparison to gene sequences which do not always represent the species under study, due in part to sequence divergence from related model plant systems (Carpentier *et al.*, 2008b). Because tef is a non-model plant system, high throughput transcriptomic approaches are somewhat limited as gene sequences are typically not conserved from one species to another, making further inferences difficult to achieve (Carpentier *et al.*, 2008b). In addition, for valid stress-responsive interpretations from previously highlighted transcripts to be drawn, a level of agreement would have to occur between the messenger transcript and translated protein (Carpentier *et al.*, 2008b). Furthermore, biological processes or actions within plant cells are performed by translated proteins rather than mRNA transcripts and transcriptomic expression does not necessarily result in protein expression (Carpentier *et al.*, 2008b). Despite these advantages, however, proteomics-based approaches are still subjected to various limitations (Chandramouli and Qian, 2009; Ow *et al*., 2009; Wasinger *et al*., 2013), some of which include: challenging sample preparation procedures (Neilson *et al*,. 2010), poor coverage of low-abundance and membrane proteins (Chandramouli and Qian, 2009; Wasinger *et al*., 2013), accurate quantitation methods that avoid intrinsic noise (Ong and Mann, 2005) and the occurrence of low fold changes (low orders of magnitude) during expressional regulation (Ow *et al*., 2009).

1.5.1 Mass spectrometry-based approaches

In previous years, the most common method of protein analysis involved ‘gel-based’ methods where proteins were separated in two dimensions (two-dimensional electrophoresis, 2-DE) according to molecular mass and isoelectric points (O'Farrell, 1975), followed by staining methods with various dyes to identify and visualise proteins of interest (two-dimensional difference gel electrophoresis, 2D-DIGE) (Unlu *et al.*, 1997; Baginsky, 2009; Nanjo *et al.*, 2011; Agrawal *et al.*, 2013). This ‘gel-based’ approach, although still commonly used today as a reliable method of protein profile visualisation, is gradually being replaced by ‘gel-free’ systems, through shotgun proteomic methods, usually involving fractionation techniques coupled to high accuracy mass spectrometry (MS) (Vanderschuren *et al.*, 2013). The shift from ‘gel-based’ systems to mass spectrometry-based proteomic approaches is mostly due to poor reproducibility occurring in the former, with increased sensitivity, extended dynamic range and overall

better proteome coverage occurring in the latter (Nanjo *et al.*, 2011; Abreu *et al.*, 2013; Vanderschuren *et al.*, 2013).

Among the most widely used gel-free quantitative proteomic approaches is iTRAQ (Isobaric Tag for Relative and Absolute Quantitation) (Ross *et al.*, 2004; Jorrin-Novo *et al.*, 2009; Nanjo *et al.*, 2011; Abreu *et al.*, 2013). iTRAQ allows digested peptides from separate samples to be differentially labelled with chemically identical tags that differ in mass only and are combined for analysis during a single tandem mass spectrometry (MS/MS) run (Jorrin-Novo *et al.*, 2009; Mochida and Shinozaki, 2010). The peptides are labelled either at the N-terminus or at lysine residues and the MS/MS allows specific fragmentation of the tag. The difference in mass between the tags allows their intensity to be used in the relative comparison of the abundance of the two or more samples, allowing quantitative information to be inferred, while fragmentation of the peptide allows protein identity to be obtained (Jorrin-Novo *et al.*, 2009; Mochida and Shinozaki, 2010).

The use of quantitative proteomic methods such as these have become a powerful and widely-used technique in the field of crop stress tolerance research as it has the ability to identify and quantify changing stress-related proteins (Baginsky, 2009; Barkla *et al.*, 2013). For comparative proteomic studies, particularly between stress-treated and untreated experiments or for further insight into the proteomic profiles of stress-sensitive to stress-tolerant crops, high-throughput, quantitative proteomics has become especially valuable (Nanjo *et al.*, 2011). To accompany protein detection and quantification by high throughput proteomic approaches, using the most comprehensive database is of equal importance for protein identification and for use in further downstream bioinformatics analyses (Cañas *et al.*, 2006; Neilson *et al.*, 2010; Balbuena *et al.*, 2011; Nanjo *et al.*, 2011). With the current availability of the sequenced tef genome and transcriptome, a large amount of accurate conclusions can be drawn from the use of proteomic methods to provide further insight in tef stress response, particularly to confer and elaborate on drought tolerance mechanisms.

1.5.2 Proteomics to study stress response in crop plants

Proteomic studies have led to the discovery of a number of stress-related proteins and have facilitated attempts to explore their importance in improving plant yield and tolerance to various environmental stresses (Salekdeh and Komatsu, 2007; Mochida and Shinozaki, 2010; Benkeblia, 2011). In addition, the iTRAQ method of identifying and quantifying protein abundance changes have been used in a multitude of proteomic analyses in commercially important crops (Nanjo *et al.*, 2011). For example, iTRAQ analysis has been used in investigations of: soybean for improved cultivar development (Qin *et al.*, 2013), into the development, metabolism and ripening of grape berry (Kambiranda *et al.*, 2013; Martinez-Esteso *et al.*, 2013), in observing the effect of storage in harvested cassava roots (Owiti *et al.*, 2011) and for investigations of the proteins causing grain chalkiness in rice (Lin *et al.*, 2014).

The available literature pertaining to proteomic investigations in crop plants subjected to various stress conditions, such as heat, cold, drought, salinity, water-logging and in response to heavy metals, are extensive and many reviews have detailed their progress thus far (see Salekdeh and Komatsu, 2007; Ahsan *et al.*, 2009; Hashiguchi *et al.*, 2010; Neilson *et al.*, 2010; Kosová *et al.*, 2011; Nanjo *et al.*, 2011; Agrawal *et al.*, 2013; Barkla *et al.*, 2013). A few quantitative proteomic studies performed in agriculturally-important crops subjected to stress conditions include investigations in: rice exposed to various stresses (Agrawal and Rakwal, 2006; Chitteti and Peng, 2007; Kim *et al.*, 2014; Wang *et al.*, 2014), soybean in response to water-logging (Alam *et al.*, 2010), wheat in response to increased salinity (Fercha *et al.*, 2014), sorghum in response to heat and drought stress (Johnson *et al.*, 2014), grape and grapevine leaves in response to drought and increased salinity (Vincent *et al.*, 2007; Liu *et al.*, 2014) and barley to observe tolerance to the heavy metal, boron (Patterson *et al.*, 2007).

Furthermore, proteomic methods have been used to identify proteins active in response to drought stress in crops, such as: maize (Benešová *et al.*, 2012; Yang *et al.*, 2014), wheat (Ford *et al.*, 2011; Jiang *et al.*, 2012; Budak *et al.*, 2013), sorghum (Jedmowski *et al.*, 2014), chickpea (Pandey *et al*., 2008; Kumar *et al*., 2014) and rice (Salekdeh *et al.*, 2002; Dong *et al.*, 2014). In many of these studies various insights into the proteins changing during seedling germination, plant growth, development and their consequential effects on subsequent grain yields under drought stress conditions have been expanded. In addition to stress-related protein identification, stress responsive pathways have been highlighted (Hashiguchi *et al.*, 2010; Kosová *et al.*, 2011). These would include plant maintenance and metabolism pathways, particularly carbon and energy metabolism (proteins associated with photosynthetic and electron transport reactions), carbohydrate metabolism (proteins active in glycolysis, and biosynthesis of sugars and other oligosaccharides) as well as pathways active in early stress detection (stress-inducible signalling pathways, reactive oxygen species (ROS), ROS acting as stress signals etc.) and stress acclimation (stress protective proteins such as chaperones, ROS scavenging enzymes, late embryogenesis abundant "LEA" proteins) (Hashiguchi *et al.*, 2010; Kosová *et al.*, 2011).

1.5.3 Proteomics research in tef

To date, there has been no published proteomic study on tef with respect to an in-depth protein profiling or comparative proteomics study. The previous protein studies that have been conducted on tef were mostly targeted to the amino acid composition of tef seeds (Lester and Bekele, 1981) and the characterisation of albumin, globulin and prolamin contents in relation to nutritional quality during tef grain consumption (Bekele *et al.*, 1995; Tatham *et al.*, 1996). Tef, however, has not yet been subjected to a high-throughput, comprehensive proteomic investigation, because most studies have focused on enhancing tef productivity using genetic, genomic and cross hybridisation methods.

1.6 Objectives of this study

The objectives of this study were firstly, to physiologically characterise the response of pre-flowering tef plants to controlled dehydration stress conditions. Changes in water content were monitored using both relative water content (RWC) and absolute water content (AWC) analysis, while membrane permeability and photosynthetic potential under applied stress conditions, were monitored using electrolyte leakage measurements and chlorophyll fluorescence analysis, respectively. Photosynthesis is a metabolism required for on-going growth of plants, it is also highly sensitive to water-deficit stress and has been cited as one of the major causes of viability loss under drought conditions (Smirnoff, 1993; Foyer and Noctor, 2009). Electrolyte leakage gives a measure of membrane integrity and thus plant viability. Ultra-structural analysis was used to observe the subcellular organisation and changes therein of tef leaf tissues during dehydration stress, as well as for confirmation of other physiological parameters assessed.

The second objective was to conduct an in-depth proteomic analyses in leaf tissues of tef by identifying and quantifying differentially expressed total proteins changing in abundance levels under dehydration stress conditions. This was achieved through the use of iTRAQ mass spectrometry and appropriate database searching. As part of a comprehensive protein study, various bioinformatics tools were employed to observe and further characterise stress responsive proteins, either changing on their own or in concert with a suite of proteins in a particular pathway to stress conditions. Lastly, the third objective of this study was to biologically validate a subset of differentially regulated proteins to confirm the iTRAQ protein results and infer biological relevance to stress conditions. Protein abundance changes were observed through immunodetection by western blotting, while the relative activities of enzymes chosen for validation purposes were measured using appropriate assays.

This study, to our knowledge, is the first proteomic analysis of tef in response to water-deficit stress.

Chapter 2: Tef physiological characterisation

2.1 Introduction

When abiotic stress conditions such as salinity, drought and extreme temperatures occur in plants, a wide range of physiological, biochemical and molecular processes are modified (Ashraf and Harris, 2013). Among the stresses, drought stress has been proposed to be the most important abiotic factor in limiting crop plant growth, development and productivity (Reddy *et al.*, 2004; Benešová *et al.*, 2012; Takele and Farrant, 2013). With the increase in environmental drought conditions and the consequential need for crop plants with improved drought tolerance, the study of the molecular mechanisms induced by drought is critical to understanding whole-plant responses to stress conditions.

A significant indicator of subcellular damage associated with water deficit imposed by drought stress, is electrolyte leakage (Bajji *et al.*, 2002; Molaei *et al.*, 2012). This measure allows the determination of cell membrane permeability and damage caused by stress conditions and is generally accompanied by increased rates of electrolyte leakage over unstressed tissues (Blum and Ebercon, 1981; Kocheva *et al.*, 2004; Molaei *et al.*, 2012). The leakage rate of solutes and cell components from plant cells as a consequence of dehydration stress has been established as an accurate measure of membrane integrity and has been used to assess damage to plant cells during environmental stress conditions (Bewley, 1979; Premachandra and Shimada, 1987; Kocheva *et al.*, 2004). Thus, the observation of electrolyte leakage rates could be used to infer critical periods of stress during which subcellular damage occurs.

Photosynthesis is a key process contributing to plant growth, maintenance and development and is extremely vulnerable to changes in environmental conditions (Reddy *et al.*, 2004; Ashraf and Harris, 2013). The disruption of photosynthesis as a consequence of water deficit is one of the main causes of oxidative stress due to excessive ROS production (Smirnoff, 1993; Foyer and Noctor, 2009). The efficiency of photosynthesis can be observed in plants undergoing stress using chlorophyll fluorescence, where the total quantum yield (F_v/F_m) and quantum yield of photosystem II (PSII) (ϕ_{PSII}) is measured in real-time (Kocheva *et al.*, 2004). The use of electron microscopy to examine cell ultrastructure of plant tissues during different stages of water-deficit stress provides valuable information and insight in how plant cell walls, membranes and various organelles such as vacuoles, nuclei and chloroplasts are affected by the stress (Sherwin and Farrant, 1996; Farrant, 2000).

Previously, various physiological tests in the form of changing cell membrane stability, osmotic adjustment, gas exchange measurements, stomatal conductance and ultra-structural studies have been conducted in tef in response to water loss (Shiferaw and Baker, 1996; Ayele *et al.*, 2001; Degu *et al.*, 2008; Ginbot and Farrant, 2011). While numerous inferences have been gained from these studies, these have been conducted on different tef varieties and on plants of varying developmental stages.

Ginbot and Farrant (2011) have previously shown critical water loss stages accompanied by physiological testing in four-week-old tef plants. To elaborate on these critical water-loss stages and previous physiological findings in younger plants as well as to observe if drought-tolerance mechanisms are uniform or change with plant development, in the current study seven-week-old, mature tef plants were subjected to controlled dehydration treatment accompanied by physiological measurements. The reasoning behind testing at this stage is that it is the stage just prior to flowering, where energy shifts in preparation for inflorescence structures and seed development, which will ultimately determine crop yield. Because the largest concern of environmental drought impacts tef plant growth and grain yield productivity (Degu *et al.*, 2008; Degu and Fujimura, 2010), conducting physiological testing and subsequent proteomic analyses under drought-simulated conditions at stages just before flowering and seed development would provide insight into whole plant drought-tolerance mechanisms in an agricultural context.

In this chapter, the changing physiological parameters of approximately seven-week-old tef plants under imposed controlled dehydration stress were investigated. These parameters included measuring water loss from leaf tissues using relative water content (RWC) and absolute water content (AWC) analyses, measuring electrolyte leakage as an indicator of cell membrane integrity, measuring chlorophyll fluorescence to determine photosynthetic potential and the use of transmission electron microscopy (TEM) to visually assess damages associated with water loss in tef plant cells.

2.2 Materials and Methods

2.2.1 Tef plant germination and growth conditions

Tef plantlets were germinated from seed (brown-seeded, local market variety purchased in Ethiopia) into 6 trays (length = 30 cm, width = 27 cm and depth = 11 cm) each containing 4 kg soil mix (2.5 parts potting soil, 2 parts peat vermiculite mix (Sunshine mix 1, SunGro Horticulture) and 1 part quartz sand). The trays were then soaked with water and left to imbibe and settle for 15 min before sprinkling a few seeds onto the top layer of soil. To ensure minimum disruption, the seeds were sprayed with water using a spray canister until well moistened, followed by an additional spray with 0.114% (w/v) phostrogen (NPK: 14:10:27 and trace elements, Bayer) to further aid seed germination. The trays were then covered with plastic wrap to prevent moisture loss and left to germinate under plant growth room conditions (16 h light and 8 h dark, temperature of 25 °C, relative humidity of 45-50% and light intensities ranging from 135-150 µmol. m^{-2}. sec^{-1}) for 1 week before the plastic wrap was removed. Following 1 week of germination, tef plantlets were watered twice weekly to allow adequate plant growth and development for at least 6 weeks before imposing dehydration stress. Furthermore, tef plants were fertilised twice with 0.114% (w/v) phostrogen during the plant growth period before initiating stress treatment.

2.2.2 Tef dehydration stress treatment

2.2.2.1 Tef dehydration stress and sampling

Approximately 10 days prior to dehydration treatment, six-week-old tef plants were moved to a plant growth chamber (Percival Intellus control system, model number: I-41LL) and incubated under controlled conditions of 25 °C, 14 h day with light intensities of approximately 153-163 µmol. m^{-2}. sec^{-1}; 17 °C, 10 h night. During a 10-day acclimation period, tef plants were watered every 2 days with 500 ml water. Subsequent to acclimation, dehydration stress was imposed by withholding water for a period of 20 days from 3 trays of tef plants designated D1 to D3 (dehydrated experimental biological repeats, one to three for plants subjected to dehydration treatment), while the remaining 3 trays designated H1 to H3 (hydrated control biological repeats, one to three for plants maintained in the hydrated state) were maintained with 500 ml water every 2nd day. Because tef leaves are thin and narrow, it was established that each tray of tef plants would act as a biological repeat of pooled plants in order to have enough leaf material for all further testing.

During the dehydration period, leaves were sampled for water content analysis by gravimetric methods and used to calculate both absolute water content (AWC) to observe the weight of water lost and relative water content (RWC) to observe the loss of water relative to fully hydrated, turgid conditions. Soil moisture measurements were determined using a soil probe (HH2 Moisture Meter, Delta-T Devices, UK). Plants were sampled on the day at which water was withheld to establish starting water contents and then again on day 5 once a substantial drop in soil moisture was observed, after which sampling occurred every 2 to 3 days until the end of the dehydration period. Following the 20th day of dehydration treatment

of tef plants, the dehydrated trays (D1-D3) were rehydrated with 500 ml water to observe tef plant recovery. At each sampling time point during dehydration, tef leaf material (approximately 15 leaves, randomly selected) were flash frozen in liquid nitrogen and stored at -80 °C until further use in total protein extractions and biological validation procedures.

2.2.2.2 Absolute water content (AWC) and relative water content (RWC) determination

Leaf RWC and AWC measurements were determined according to Farrant (2000) with a few modifications. A total of six leaves were randomly sampled from each of the three trays undergoing dehydration treatment (D1-D3) and each of the three trays acting as a hydrated control (H1-H3), following the standard formula:

AWC ($gH_2O.gdw^{-1}$) = (fresh weight – dry weight)/ dry weight,

RWC (%) = ($AWC_{sample}/AWC_{full\ turgor}$) x 100

The leaves were immediately weighed using an ultra-fine balance (Mettler Toledo, USA) after sampling to obtain fresh weights, followed by floating leaves in water for 24 h to allow for maximal water uptake and thus obtain full turgor weight for use in $AWC_{full\ turgor}$ measurements. Thus for full turgor weight measurements ($AWC_{full\ turgor}$) and subsequent RWC calculations, each leaf was made relative to itself at fully hydrated, turgor conditions. Prior to weighing leaves for full turgor weight, leaves were gently wiped with paper towel to remove excess water. Dry weights were also gravimetrically determined by oven drying at 70 °C for 48 h, followed by cooling in a desiccator for 10 min before weighing leaf samples again to obtain dry weight values.

2.2.3 Electrolyte leakage

The rate of electrolyte leakage from leaves of tef plants during dehydration stress was measured using a CM 100-2 Multiple Cell Conductivity Meter (Reid & Associates, South Africa). A total of three tef leaves were sampled from dehydrated (D1-D3) replicate trays and cut into 1.5 cm long segments. Leaf segments were equally distributed into plastic well trays (approximately 2 ml in volume). Thereafter, 1.5 ml ultrapure water was added to the wells and conductivity measurements were immediately started. Measurements were taken every 1 min over a 20 min period. Subsequent to conductivity measurements, leaf samples in each of the wells were placed into labelled foil packets and dried by oven drying at 70 °C for 48 h, before being placed in a desiccator for 10 min and weighed to obtain the dry weights. The rate of electrolyte leakage was calculated by plotting electrolyte leakage values on a straight line to obtain the gradient of the line (rate of change over time for increasing values only) and used in the following equation:

Electrolyte leakage = rate of leakage/dry weight of leaf segments, where the rate of leakage was expressed as $\mu S.\ min^{-1}.gdw^{-1}$.

2.2.4 Chlorophyll fluorescence

Chlorophyll fluorescence measurement was performed according to Maxwell and Johnson (2000), using a portable PAM-2100 Chlorophyll fluorometer (Walz, Germany). Approximately three tef leaves were aligned in order to cover the area of the dark adaption clips (4 mm in diameter). Leaves were dark-adapted for 15 min before maximum quantum yield of PS II (F_v/F_m) values were calculated using the standard formula:

$$F_v/F_m = (F_m-F_0)/F_m$$

(where F_m is the maximum fluorescence yield of PS II after a saturating light pulse and F_0 is the baseline fluorescence of dark adapted leaves).

In addition the quantum yield of PS II (ϕ_{PSII}) was also calculated from the formula:

$$\phi_{PSII} = (F'_m-F_t)/ F'_m$$

(where F'_m is the fluorescence maximum of light-adapted leaves and F_t is the steady-state value of fluorescence just before initiation of saturating light pulse and actinic light. Chlorophyll fluorescence measurements were performed in triplicate on each of the tef biological trays under dehydrated conditions (D1-D3).

2.2.5 Transmission Electron Microscopy (TEM)

In order to investigate cellular ultrastructure of tef leaves undergoing dehydration treatment in leaf tissues where plant cells were most viable, a section of leaf tissue slightly above the basal meristematic regions were examined using transmission electron microscopy (TEM).

2.2.5.1 Chemical fixation

Tissue preparation was carried out according to the method described by Sherwin and Farrant (1996), with a few modifications. For chemical fixation, a section of leaf tissue was excised approximately 1 cm away from the base of the leaf. The section was then cut up into smaller segments (1-2 mm wide) with a sharp blade and incubated in gluteraldehyde solution (2.5% (v/v) gluteraldehyde, 0.1 M phosphate buffer, pH 7.4, 0.5% (w/v) caffeine) overnight at 4 °C. After incubation at 4 °C, leaf segments were washed three times with 0.1 M phosphate buffer, pH 7.4, for 5 min each before being post-fixed in osmium tetraoxide (1% (v/v) OsO_4 (Sigma-Aldrich, Inc.) in 0.1 M phosphate buffer, pH 7.4) for 1 h. Subsequent to fixation, leaf segments were subjected to three washes in 0.1 M phosphate buffer, pH 7.4, for 5 min each. The samples were then dehydrated in a graded ethanol series of 30%, 50%, 70%, 90% and 100% (v/v) ethanol, by incubating leaf segments twice for 5 min in each ethanol concentration. The leaf segments were then incubated twice in 100% (v/v) acetone, for 10 min each before adding an equal volume (to that of acetone) of Spurr's resin (Spurr, 1969), followed by incubation overnight at 4 °C. Gradually, the acetone was replaced with increasing amounts of resin until samples were in 100% pure resin, incubating overnight at 4 °C at each concentration. Samples were polymerised at 60 °C for 16 h.

2.2.5.2 Staining and electron microscope viewing

Embedded samples were sectioned at 95 nm using a Diatome diamond knife (Diatome, Switzerland) on a Reichert Ultracut S Ultra-microtome (Leica, Austria) and mounted onto copper grids. The sections were stained with 2% (w/v) uranyl acetate for 10 min. The sections were washed 5 times with ultrapure water for 20 sec per wash, before being stained with Reynolds lead citrate (Reynolds, 1963) for 10 min. The sections were then washed one more time by jet-washing of grids with ultrapure water before blotting on filter paper and viewed with a FEI/Tecnai T20 (FEI, USA) microscope.

2.3 Results and Discussion

2.3.1 Tef dehydration treatment

Tef leaves maintained a water content of approximately 80-90% RWC (~3 $gH_2O.gdw^{-1}$) for six days before a gradual loss of water was observed, over a period of 20 days, leaves lost *ca.* 60-70% (~2 $gH_2O.gdw^{-1}$) of their internal water (Fig. 2.1A, B) but recovery was observed in plants that had lost such water when re-watered (data not shown). During dehydration, at 13 days, water content was approximately 50% RWC (~1.5 $gH_2O.gdw^{-1}$) and plants showed signs of leaf curling and stems had yellowed (Fig. 2.1D). Leaf folding and rolling is a common morphological change, particularly among cereals, undergoing abiotic and biotic stress (reviewed in Kadioglu *et al.*, 2012). The curling or rolling of the leaves reduces the leaf surface area and thereby results in benefitting the plants in two ways. Firstly, it causes reduced light exposure resulting in less radiation damage and minimises photosynthesis-induced oxidative stress (Sarieva *et al.*, 2010). Secondly, it increases humidity close to the leaf surface, which results in reduced transpiration rate (Tanimoto and Itoh, 2001).

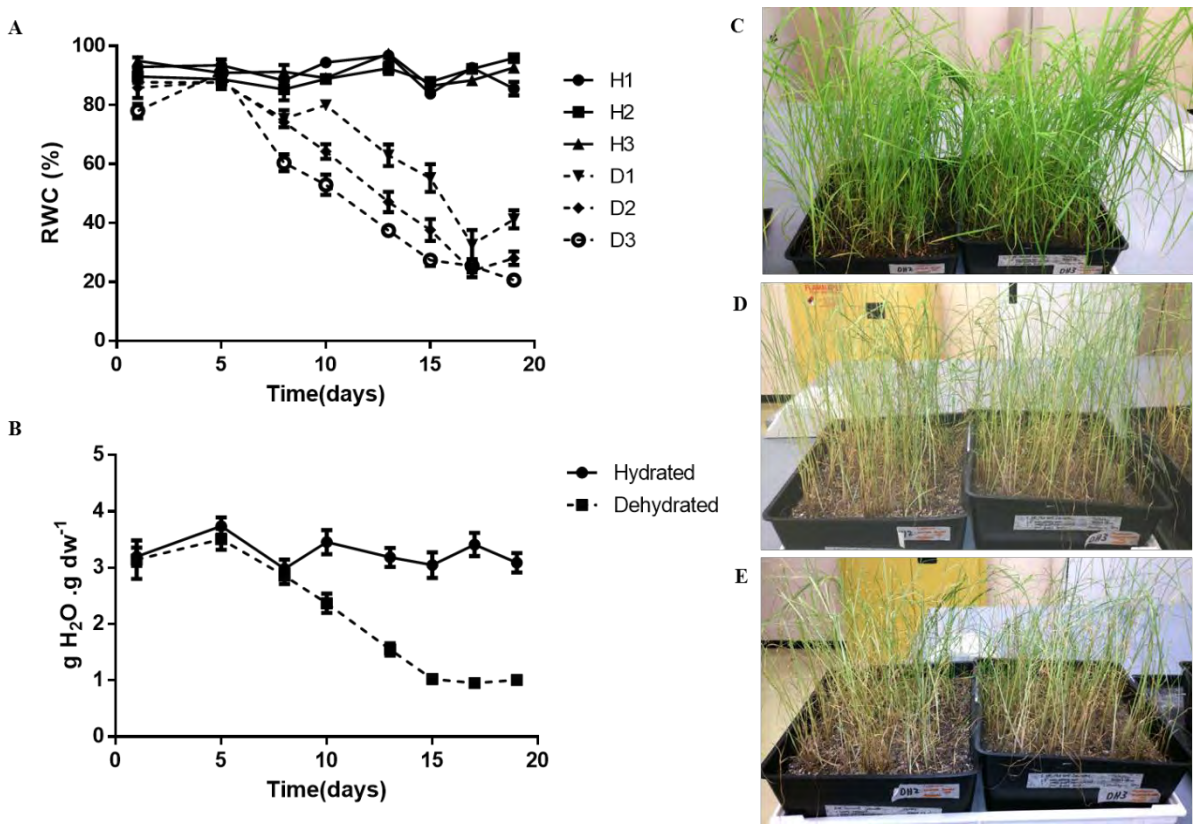

Fig. 2.1 Relative water content (RWC) curve (A) and absolute water content (AWC) curve (B) of tef plants subjected to dehydration treatment. Hydrated tef plants (C) at the start of dehydration (day 1, ~85% RWC, ~3 $gH_2O.gdw^{-1}$), dehydrated tef plants after 13 days of no water at approximately 50% RWC, ~1.5 $gH_2O.gdw^{-1}$ (D) and dehydrated tef after 17 days of no water at approximately 25% RWC, ~1 $gH_2O.gdw^{-1}$ (E). Solid lines in curves denote hydrated plants (control, H1- H3) and dashed lines denote dehydrated plants (experimental, D1- D3), where RWC values shown are means of 5 replicates (n = 5) (A). AWC values shown are the means of pooled replicates for both hydrated and dehydrated curves at set time points ($n \geq 10$) (B). Error bars represent standard error between replicates.

After 17 days of no water, leaf water contents of approximately 25% RWC (~1 $gH_2O.gdw^{-1}$) were reached and plants showed clear signs of dehydration stress-induced injury where leaves appeared shrivelled and wilted (Fig. 2.1E). At RWCs below 20% RWC (< 1 $gH_2O.gdw^{-1}$), lack of recovery on re-watering indicated that there was irrevocable damage and plant cell death had occurred.

2.3.2 Electrolyte leakage

To investigate the permeability of membranes and efflux of electrolytes from plant cells with dehydration stress, tef leaves dried to various RWCs were subjected to conductivity measurements as shown in Figure 2.2. Upon dehydration, a progressive increase in the rate of electrolytes lost was observed for dehydrated tef leaves to a value of 570 $\mu S.min^{-1}.gdw^{-1}$ (Fig. 2.2). The increase in electrolyte leakage for tef leaf samples at the early stages of dehydration treatment (80-90% RWC) (Fig. 2.2), could be due to a consequence of cutting leaves in preparation for conductivity measurements. However, because cut surfaces were uniform among treatments, it was assumed that increased leakage above this initial value (160 $\mu S.min^{-1}.gdw^{-1}$) would indicate water-deficit induced membrane damage. A slight decline in electrolyte leakage was observed between 40-55% RWC before rapidly increasing to a maximum rate of 780 $\mu S.min^{-1}.gdw^{-1}$ (Fig. 2.2). The decrease in electrolyte leakage below 20% RWC could potentially be due to the measurement of dead leaf tissue.

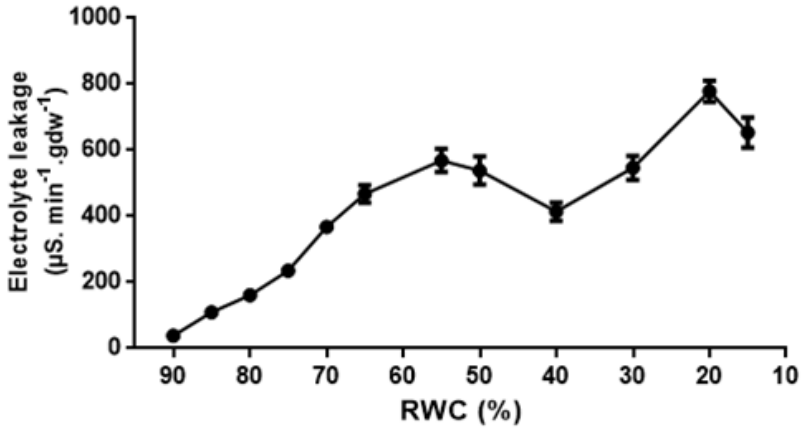

Fig. 2.2 Electrolyte leakage of tef plants subjected to dehydration treatment in a decreasing RWC range of 90 to 10% RWC. Values shown are means of 6 replicates ($n \geq 6$) pooled from dehydrated plants at designated RWCs. Error bars denote standard error.

The plasma membrane has been reported to be the primary site of structural damage during dehydration stress (Levitt, 1980; Bajji *et al.*, 2002; Molaei *et al.*, 2012). In the initial stages of dehydration, a rapid increase in electrolyte leakage usually occurs from free intracellular spaces (Bajji *et al.*, 2002), as seen from the electrolyte leakage curve until 55% RWC (Fig. 2.2). Because the overall trend in electrolyte leakage seems to increase with continuous dehydration treatment, the decline in electrolyte leakage activity observed below 55% RWC (Fig. 2.2) may perhaps be due to drought-tolerance protection mechanisms occurring in tef, in an attempt to minimise cellular water loss. Alternatively, an early

increase in electrolyte leakage can occur in some plants undergoing certain stresses without membrane damage (Bajji *et al.*, 2002; Rolny *et al.*, 2011). Bajji *et al.* (2002) suggested that an increase in organic ions and not membrane damage was the reason for the increase in electrolyte leakage in Durum wheat undergoing osmotic stress. Rolny *et al.* (2011) showed that dark-induced senesced barley leaves still maintained membrane integrity and the apparent increase in electrolyte leakage could have been due to ammonium accumulation as a result of the breakdown of chloroplasts in mesophyll cells.

The loss of electrolytes during the latter stages of dehydration, are reported to be indicative of electrolytes released through the plasma membrane and vacuole tonoplast (Bajji *et al.*, 2002). As seen in Figure 2.2, below 40% RWC, irrevocable damage appears to have occurred to plant cell membranes resulting in the release of cellular constituents into the external environment and a maximum electrolyte leakage rate of approximately 780 $\mu S.min^{-1}.gdw^{-1}$. Previous electrolyte leakage measurements in tef with imposed dehydration stress showed increased electrolyte leakage rates due to membrane and subcellular damage (Ginbot and Farrant, 2011). The increase in membrane permeability with continuous dehydration stress has been linked to the synthesis of ROS, which as a by-product, cause the breakdown of proteins, membrane lipids and photosynthetic pigments that function in maintaining cell membrane stability (Navari-Izzo *et al.*, 1997; Ahmadizadeh *et al.*, 2011). In addition, one can infer that due to excessive electrolyte leakage at water contents below 40% RWC, that RWC ranges just before 40% RWC (40-55% RWC) are indeed critical water content points, where drought-tolerance protection mechanisms are potentially put into place before continuous dehydration stress causes loss of cell viability.

Tef has been previously classified as moderately drought tolerant compared to other plants within the *Eragrostis* genus (Balsamo *et al.*, 2006) and studies by Ginbot and Farrant (2011) have confirmed that this species has some measure of tolerance to water-deficit under drought stress. According to Ginbot and Farrant (2011), the brown-seeded tef variety was shown to tolerate drought conditions for a longer period of time in comparison to white-seeded tef varieties until a RWC of 30% was reached. Further dehydration stress resulted in a loss of viability in tef plant cells with ultra-structural studies displaying severe dehydration-induced damage to plant cell membranes and organelles (Ginbot and Farrant, 2011). A similar result with regards to RWC analysis and electrolyte leakage was observed in this study, where after 20 days of dehydration treatment, plants decreased to below 30% RWC (~1 $gH_2O.gdw^{-1}$) (Fig. 2.1A, B). Once water contents of 20% RWC (< 1 $gH_2O.gdw^{-1}$) were reached, plants were unable to recover upon re-watering and electrolyte leakage rates shown (Fig.2.2) were at a maximum (780 $\mu S.min^{-1}.gdw^{-1}$) indicating plant cell membrane disintegration. This suggests that some varieties of tef, particularly the brown-seeded varieties (as tested here) are able to tolerate an internal water loss of 60-70% (~2 $gH_2O.gdw^{-1}$) before cell death occurs as reported by Ginbot and Farrant (2011). Nevertheless, a loss of approximately 2 gH_2O before losing viability is a substantial amount in comparison to other crop species

such as wheat (Siddique *et al.*, 2000), maize (Benešová *et al.*, 2012; Takele and Farrant, 2013) and sorghum (Takele and Farrant, 2013).

2.3.3 Chlorophyll fluorescence

To observe the effect of dehydration stress on the components of the photosynthetic machinery (photosystems and electron transport chain), chlorophyll fluorescence measurements were conducted on tef leaves subjected to dehydration stress (Fig. 2.3A, B).

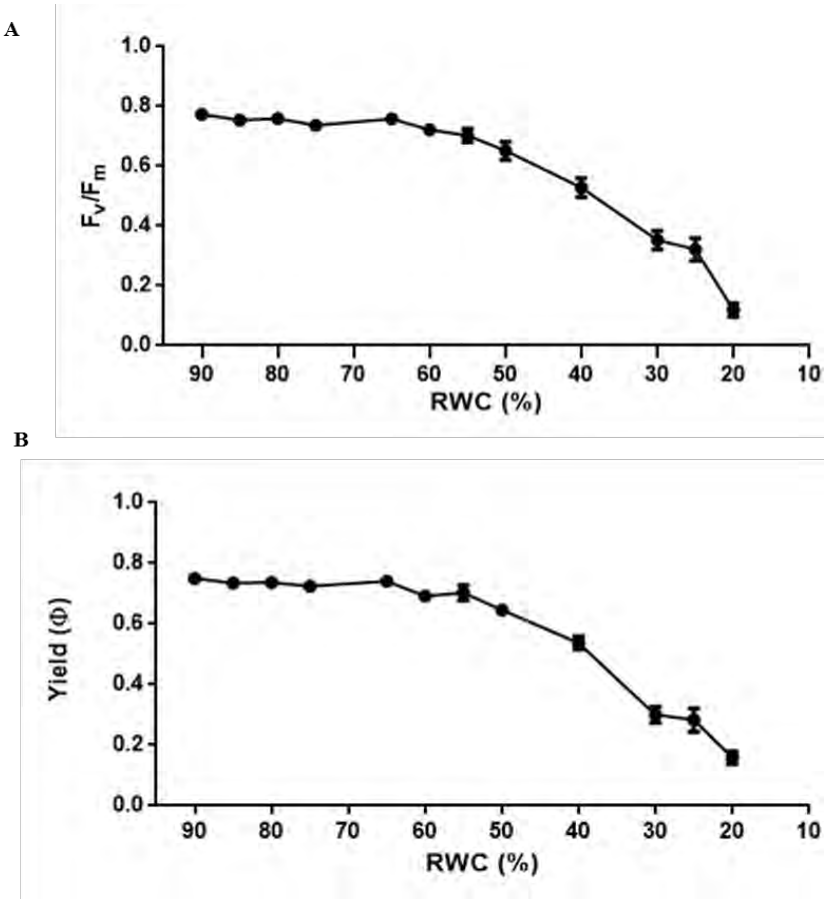

Fig. 2.3 Chlorophyll fluorescence measurements of tef plants subjected to dehydration treatment in a decreasing RWC range of 90 to 10% RWC. Maximum quantum yield of PS II (F_v/F_m) (A) and quantum yield of PS II (ϕ_{PSII}) (B). Chlorophyll fluorescence measurements were conducted in triplicate at each time point ($n \geq 3$) and error bars denote standard error between replicates.

The changes in quantum efficiency of PS II (F_v/F_m) and quantum yield of PS II (ϕ_{PSII}) with dehydration stress are shown (Fig. 2.3A, B). In both curves, a similar trend was observed, where F_v/F_m and ϕ_{PSII} were maintained at a value of approximately 0.75 until 55% RWC (Fig. 2.3A, B). An F_v/F_m value of approximately 0.75-0.85 has previously been reported to be indicative of healthy, non-stressed leaves (Jimenez *et al.*, 1997). Below 55% RWC, however, signs of disruption to PS II and electron transport occurred in tef and leaves displayed a gradual decline in photosynthetic potential (Fig. 2.3A, B).

The disruption of photosynthetic processes is potentially due to photo-oxidative damages caused by light-chlorophyll interactions and increased occurrence of free radicals as water is lost from the cell (Dace *et al.*, 1998; Farrant, 2000). The chloroplasts in particular are reported to be sensitive to photo-oxidative

damage when water is unavailable leading to proliferation of ROS (Smirnoff, 1993; Navari-Izzo *et al.*, 1997; Farrant, 2000; 2007; Ginbot and Farrant, 2011). The further decline in the quantum efficiency and yield of PS II below 55% RWC to a value of approximately 0.2 as water was gradually lost (Fig. 2.3A, B), indicated damage to thylakoid membranes (Waraich *et al.*, 2012; Ashraf and Harris, 2013) and potential loss of photosynthetic pigments leading to diminished electron transport and hence photosynthetic potential (Ashraf and Harris, 2013).

Previous studies by Ginbot and Farrant (2011) have shown that brown-seeded tef varieties have the ability to somehow minimise damage to chloroplasts and maintain photosynthetic capacity to a RWC value of 43%. The previously discussed leaf curling (leaf rolling) observed in tef with dehydration stress, as well as having a C4 photosynthetic metabolism, may be the reason for maintaining photosynthetic capacity at these RWCs. The loss of internal water below 43% RWC, however, resulted in increased damage to chloroplasts and irreversible damage to photosynthetic capacity (Ginbot and Farrant, 2011), similarly to what is observed in this study as shown in Figure 2.3A, B.

2.3.4 Ultrastructure of tef leaves during dehydration stress

TEM investigation was conducted to observe the ultra-structural changes within plant cells as dehydration treatment proceeds, displayed in Figures 2.4 to 2.6.

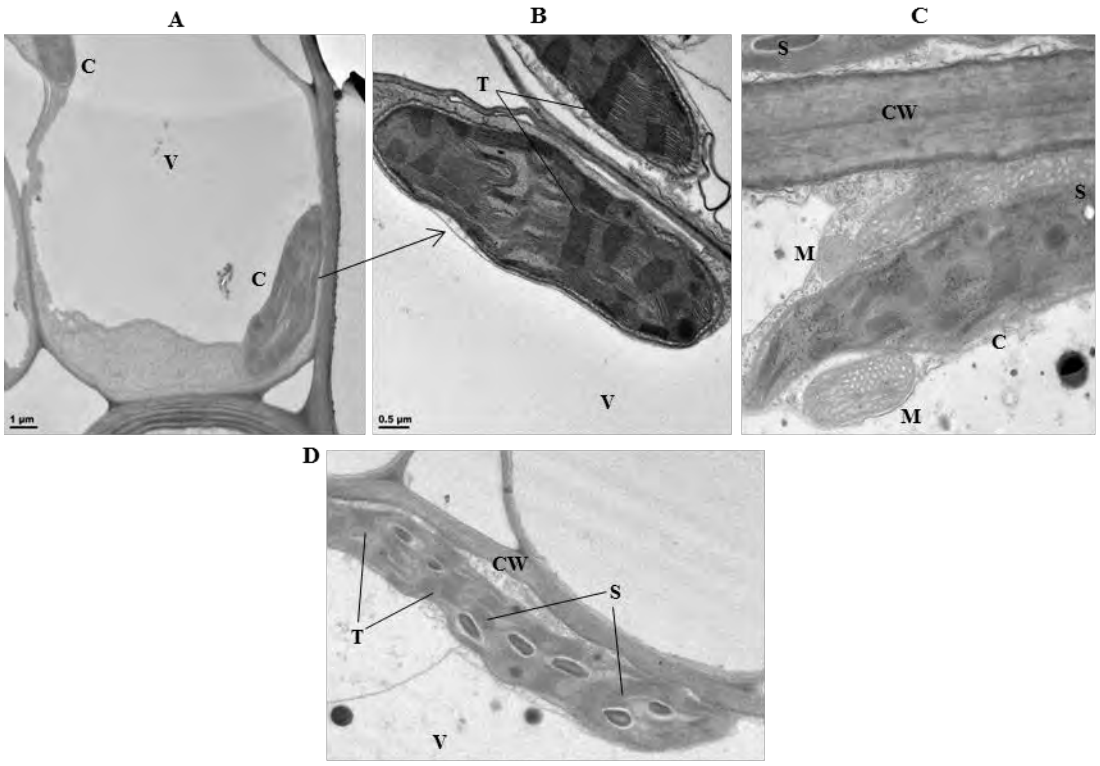

Fig. 2.4 Electron micrographs of mesophyll cells from tef leaf tissues in the hydrated state at 87% RWC (A to D). C = chloroplast, V = vacuole, T = thylakoid membranes, S = starch, CW = cell wall and M = mitochondria. Scale bar = 1µm for A and 0.5 µm for B, C and D.

Ultra-structural observations of tef leaf tissues in the hydrated state at 87% RWC (Fig. 2.4), showed healthy, non-stressed cells with clearly defined cellular components. The organelles appeared intact with a central turgid vacuole and plasma membrane appressed to the cell wall (Fig. 2.4 A to D). The chloroplasts in hydrated tissues had well-defined, stacked thylakoids (Fig. 2.4B to D) with considerable starch granules, indicating an actively photosynthesising system (Fig. 2.4C, D). Furthermore, mitochondria had well-developed, defined cristae (Fig. 2.4C), suggesting an actively respiring system.

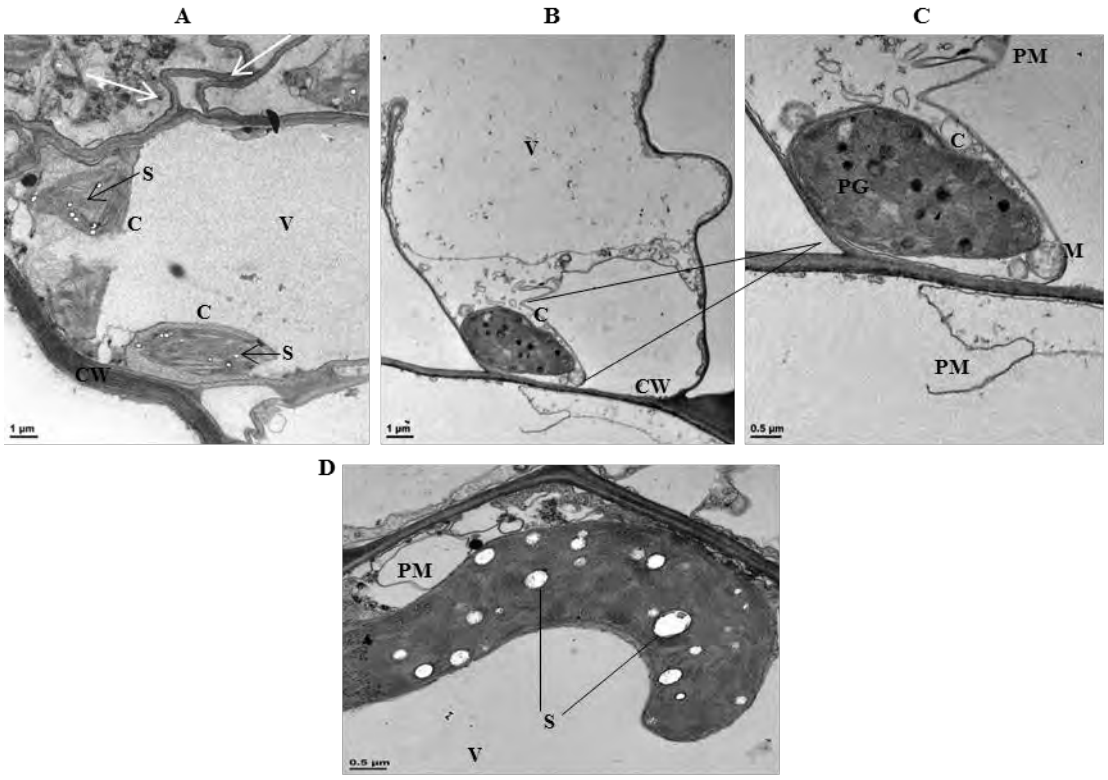

Fig. 2.5 Electron micrographs of mesophyll cells from tef leaf tissues in the dehydrated state at 60% RWC (A) and 50% RWC (B, C and D). C = chloroplast, V = vacuole, S = starch, CW = cell wall, PM = plasma membrane, M = mitochondria and PG = plastoglobuli. White arrows = cell wall folding. Scale bar = 1µm for A and B, 0.5 µm for C and D.

Dehydration to 60% RWC (Fig. 2.5A), showed a still actively metabolising cell, with slight indications of ultra-structural damage. A large vacuole was still present, however, there was evidence of cell wall folding (white arrows, Fig. 2.5A). By 50% RWC, there was some evidence of plasma membrane withdrawal from the cell wall (Fig. 2.5B to D). This could be due in part to the use of aqueous chemical fixation of tissue, which is known to partially hydrate wall tissues taking up water more rapidly than intracellular constituents. A few electron micrographs taken at this water content, however, showed signs of plasma membrane tearing (Fig. 2.5B, C), indicating ultra-structural damage.

Other evidence of damage observed at 50% RWC (Fig. 2.5B, C) were vacuole shrinkage and disruption in the stacking of thylakoid membranes, the formation of plastoglobuli with increased electron opaqueness and the lack of cristae differentiation and increased electron transparency in mitochondria (Fig. 2.5C). The formation of plastoglobuli are often indicative of light as well as desiccation stresses (Farrant *et al.*, 2003). As the plants in this study were not subject to light stress, it is likely that these structures are a consequence of photosynthetic adjustments to water loss. In addition, there were some distinct changes in the ultra-structural appearance of starch grains, where the structures increased in size and number and had a more electron opaque appearance (Fig. 2.5D). These changes were accompanied by distended thylakoids in chloroplasts at 50% RWC (Fig. 2.5 B to D), which further supports the

previous chlorophyll fluorescence findings, where a reduction in electron transport and photosynthetic capacity was evident (Fig. 2.3A, B). Although electron micrographs taken at this water content (50% RWC), showed signs of dehydration stress-induced damage and changes in the sub-cellular organisation of tef plant cells, these effects were found in a lesser proportion and to a tolerable extent, suggesting cell metabolism continued during dehydration at these water contents.

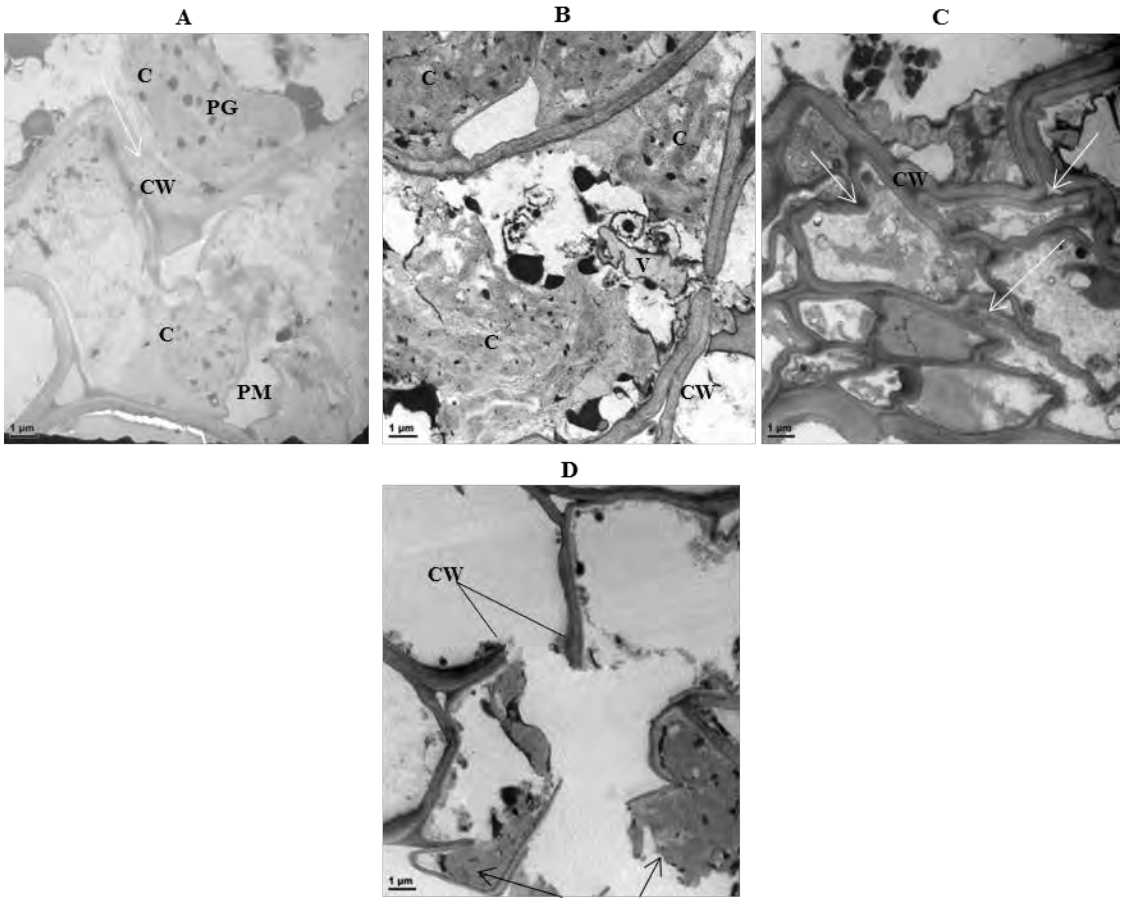

Fig. 2.6 Electron micrographs of mesophyll cells from tef leaf tissues in the dehydrated state at 35% RWC (A, B), 20% RWC (C) and less than 20% RWC (D). C = chloroplast, CW = cell wall, PG = plastoglobuli, V= vacuole and PM = plasma membrane. White arrows = cell wall folding, black arrows = cell debris. Scale bar = 1µm for A, B, C and D.

Dehydration to 35% RWC, however, showed an increase in the proportion of damaged plant cells (Fig. 2.6A, B), where cells had become compressed as liquid volume was lost, vacuole structures had decreased in size and membrane integrity was compromised. Furthermore, the increased occurrences of plastoglobuli and diminished starch granules in chloroplasts as a consequence of dehydration stress at 35% RWC (Fig. 2.6A, B) were indicative of further photosynthesis disruption. The compaction and condensation of cellular constituents with considerable evidence of cell wall folding (white arrows, Fig. 2.6A, C) and breakage (Fig 2.6D), are classical signs of dehydration stress (< 40% RWC) in desiccation sensitive tissues. At 20% RWC, largely all cells were affected by dehydration stress, where no clear definitions of cellular organelles were seen and evidence of lytic activity was observed (Fig. 2.6C). The

withdrawal and rupture of plasma membrane and damage to organelle integrity below 20% RWC (Fig. 2.6C, D), results in increased rates of electrolyte leakage as water is lost, as seen in Figure 2.2. This also confirms that the increase in electrolyte leakage during the later stages of dehydration until nearly all cellular constituents were leaked out of cells at 20% RWC (Fig. 2.2) was primarily due to membrane damage and not due to other factors.

Furthermore, the compression of cell constituents and cell wall folding shown by white arrows in Figure 2.6C, could potentially be the cause of the needle-like appearance of tef leaves (representing leaf rolling as cell volume was lost, (Fig. 2.1D, E) and change in mechanical structure of cell walls with increased dehydration stress as reported by Balsamo *et al.* (2006). In addition, the disorganised appearance of subcellular constituents and lack of compartmentalisation is similar to that reported by Ginbot and Farrant (2011). In Figure 2.6D, cell wall breakage had occurred and cellular components had been completely degraded, shown by cell debris aggregating against cell walls (black arrows, Fig. 2.6D). The presence of 'empty' cells in Figure 2.6D, at the very end of dehydration (< 20% RWC), with no definable organelles or membranes indicates that autolysis has occurred and cells have lost viability (dead tissue).

2.4 Brief conclusion

Work described in this chapter enabled characterisation of some of the physiological responses to water-deficit stress in seven-week-old (pre-flowering) tef plants. It allowed understanding of critical water contents at which damage is initiated and when this becomes damaging to plants such that viability is lost. The RWC data show that tef has the ability to retain cellular water for up to 6 days under the conditions tested, but drying over a 17 day period resulted in dehydration to approximately 30% RWC during which viability was still retained supporting previous observations by Ginbot and Farrant (2011) on the drought tolerance of this crop and of many other cereals (Blum, 1996; Takele and Farrant, 2013). It can be hypothesised that critical water loss stages occur in a range of approximately 50% RWC, as progressive changes in physiological measurements were observed just before the half-way point of dehydration.

Further imposition of dehydration stress below 30% RWC, however, results in the complete loss of water to critical water contents from which tef plants cannot recover (≤ 20% RWC). This results in membrane rupture or disintegration and loss of cellular components from plant cells with maximum electrolyte leakage rate of 780 $\mu S.min^{-1}.gdw^{-1}$ and complete photosynthetic disruption with F_v/F_m and ϕ_{PSII} values decreasing to approximately 0.2. Further ultra-structural studies have shown damage to subcellular components at water contents of 35% RWC and below, such as plasma membrane rupture towards the end of dehydration treatment, no clearly definable cellular organelles, cell wall folding and eventually breakage as well as evidence of lytic activity on cell organelles. These results further coincide with the findings previously reported by Ginbot and Farrant (2011) that some tef varieties are relatively drought tolerant, particularly the brown-seeded varieties, by having certain adaptive features that increase tolerance to drought conditions to water contents above 30% RWC. In addition, the changes in physiological measurements appear to be consistent throughout dehydration stress and coincide with what has previously been reported for younger plants at four weeks of age, indicating a level of plasticity and adaptation occurring in tef plants with dehydration stress.

Lastly, these physiological data (RWC analysis, ultra-structural analysis, electrolyte leakage and chlorophyll fluorescence measurements), have enabled selection of critical RWC stages occurring in tef with dehydration stress, for further proteomic studies conducted in the following chapters (see Chapters 3 and 4).

Chapter 3: iTRAQ analysis of tef proteins in response to dehydration stress

3.1 Introduction

One of the aims of this study was to identify and quantify tef proteins with differential expression in response to dehydration stress through the use of iTRAQ mass spectrometry and appropriate database searching. A number of physiological responses occur in tef when it is exposed to dehydration stress (see Chapter 2), such as an increase of membrane permeability (section 2.3.2) and the deterioration of photosynthetic potential (section 2.3.3), which are accompanied by changes in subcellular organisation as dehydration treatment proceeds (section 2.3.4). To observe the whole plant response to dehydration stress, it is important to compliment observations of the changing physiological parameters with an examination of changes in the total proteome.

The iTRAQ method of protein analysis was developed to quantitatively determine changes in protein abundance in biological samples. In this method, isobaric mass labels (isobaric tags) are placed at the N-termini and lysine side chains of peptides in a digested mixture, where one tag is used for each condition. Thereafter the peptides present in a sample by MS/MS scans are detected and the tags used for quantification (Ross *et al.*, 2004). By quantification of peptides directly from the mass spectra, accurate functional information as well as induced changes within the proteome can be retrieved (Ong and Mann, 2005). The iTRAQ method of identifying and quantifying proteins has been widely used in a multitude of studies (see Choe *et al.*, 2007; Wiese *et al.*, 2007; Lan *et al.*, 2011; Lin *et al.*, 2014; Martinez-Esteso *et al.*, 2014), particularly in investigating plant proteomic responses to abiotic stresses (reviewed in Nanjo *et al.*, 2011; Kosová *et al.*, 2011; Abreu *et al.*, 2013; Ghosh and Xu, 2014).
The increased usage of 'gel-free' systems such as with tandem mass spectrometry for proteomic profiling is due in part to its enhanced sensitivity, improved confidence of protein identification and reduced rate of false identification (Agrawal *et al.*, 2013; Liu *et al.*, 2014). In addition, valuable insight has been gained when iTRAQ mass spectrometry has been applied to numerous crop plants facing abiotic stress conditions including drought stress in wheat and maize (Ford *et al.*, 2011; Benešová *et al.*, 2012), temperature changes such as heat and cold stress in rice (Neilson *et al.*, 2010; 2011), and in plants exposed to heavy metals (Ahsan *et al.*, 2009).

In this present study, we have successfully used the iTRAQ technique coupled to peptide separation by OFFGEL fractionation with LC-QTOF tandem mass spectrometry analysis to observe the tef proteomic profile in response to dehydration stress. Furthermore, the use of database searching to identify proteins and statistical tools for quantification were employed to observe differential protein expression in response to dehydration stress.

3.2 Materials and Methods

3.2.1 Plant material and experimental design

Seven-week-old tef plants were dehydrated as described in chapter 2 (section 2.2.2.1, Fig. 2.1), during which leaves were sampled from hydrated (H1-H3) and dehydrated (D1-D3) trays of tef plants, for total protein extractions and further used in iTRAQ analysis described in Figure 3.1. Based on the RWC and AWC results displayed in chapter 2 (section 2.3.1, Fig. 2.1), water contents at approximately 80% RWC (~ 3 $gH_2O.gdw^{-1}$) were indicative of well-watered, hydrated tissues. Further physiological testing with dehydration stress showed the measured physiological parameters (membrane permeability, photosynthetic potential and ultra-structural analysis in chapter 2, sections 2.3.2 to 2.3.4) were changed in a RWC range of approximately 50% (~ 1.5 $gH_2O.gdw^{-1}$). For these reasons, six RWC points (three hydrated and three dehydrated) were chosen for proteomic analysis to represent proteins changing in quantitative expression during hydrated (control) and dehydrated (experimental) conditions in a range of RWCs at approximately 80 and 50% RWC, respectively. For the hydrated protein samples, leaf tissues were chosen at the water contents 88, 85 and 80% RWC and designated H1, H2 and H3, respectively. Similarly, for the dehydrated protein samples, leaf tissues were chosen at the water contents 58, 54 and 52% RWC and designated D1, D2 and D3, respectively.

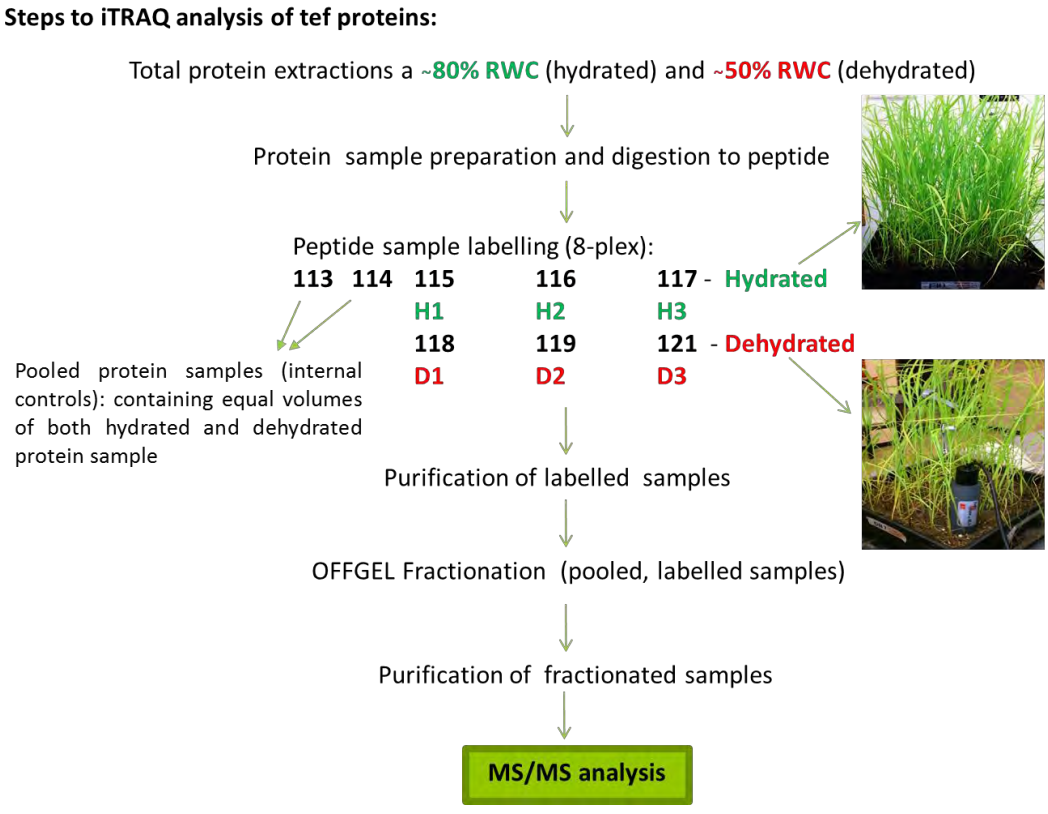

Fig. 3.1 iTRAQ analysis workflow of tef hydrated (H1-H3) and dehydrated (D1-D3) biological protein samples in preparation for mass spectrometry analysis. An 8-plex iTRAQ experimental system was used. Hydrated protein samples (H1-H3) were represented by iTRAQ labels 115-117, respectively, while dehydrated protein samples (D1-D3) were represented by iTRAQ labels 118-121, respectively. Labels 113 and 114 were used as internal control standards containing equal amounts of both hydrated and dehydrated pooled protein.

To provide biological replicates to increase the reliability of the results while maintaining affordable costs, one 8-plex iTRAQ experimental system was used with three biological replicates for each treatment, hydrated and dehydrated as well as two labels used for internal controls (Fig. 3.1). Three biological replicates at hydrated conditions of approximately 80% RWC (88, 85 and 80% RWC;~ 3 $gH_2O.gdw^{-1}$) and three biological replicates at dehydrated conditions of approximately 50% RWC (58, 54 and 52% RWC;~1.5 $gH_2O.gdw^{-1}$), were subjected to iTRAQ analysis.

3.2.2 Protein extraction

Total leaf proteins from hydrated and dehydrated tef tissues were extracted for each RWC point (a total of six extractions) according to the method by Isaacson *et al.* (2006) with a few modifications. Leaf tissue was ground in liquid nitrogen with a chilled mortar and pestle to a fine powder with the addition of 1% (w/w) insoluble polyvinylpolypyrrolidone (PVPP). Ground tissue was then aliquoted into 2 ml centrifuge tubes up to 0.1 ml mark and 1 ml ice-cold extraction buffer (0.7 M sucrose, 0.1 M KCl, 0.5 M Tris-HCl, pH 7.5, and 50 mM EDTA) together with 1 ml Tris (0.5 M, pH 8.0)-saturated phenol was added. A protease inhibitor tablet (1 Roche Complete Mini tablet per 45ml volume of extraction buffer) and the reducing reagent dithiothreitol (DTT) as well as serine protease inhibitor phenylmethylsulfonyl fluoride (PMSF) was added to the extraction buffer at a final concentration of 2% (w/v) and 1 mM respectively, just before use. The samples were then well-mixed by vortexing for 15 min at 4 °C, followed by centrifugation at 12,000 *x g* for 10 min at 4 °C to allow phase separation. Once centrifugation was complete, the upper phenolic phase containing phenol soluble proteins was carefully removed (without disturbing the white inter-phase) and transferred to a new centrifuge tube while the lower aqueous phase containing all cell debris and contaminants was discarded. An equal volume of fresh extraction buffer to that of the collected phenolic phase was added and the mixture was vortexed for 10 min at 4 °C. The samples were once again centrifuged at 12,000 *x g* for 10 min to recover the protein containing phenolic phase. To precipitate the proteins, 5 volumes (to that of the collected phenolic phase) of cold 0.1 M ammonium acetate in methanol was added and samples were incubated at -20 °C for 16 h or until further use in the filter assisted sample preparation procedure (FASP).

3.2.3 Protein quantification

Protein pellets were recovered by centrifugation of incubated samples at 12,000 *x g* for 15 min at 4 °C and the resulting supernatant was removed and discarded. The protein pellets were washed once with 1 ml 100% methanol to remove phenol, ammonium acetate, lipids and pigments at 12,000 *x g* for 5 min at 4 °C, followed by an additional wash with 80% (v/v) acetone at 4 °C to remove traces of methanol and to allow rapid drying. Protein pellets were then air-dried under a fume hood for 5 min, followed by re-suspension in 70-100 µl of 2% (w/v) SDS and vortexing for 15 min at room temperature for protein re-solubilisation. Additionally, samples were placed on a heating block at 90 °C for 3-5 min to facilitate dissolving of the pellet. The re-suspended proteins were quantified using the Pierce BCA protein assay

kit (Thermo Fisher Scientific, Inc., USA) according to the manufacturer's instructions using BSA as a standard. Protein content of samples was measured at 595 nm using the Thermo Fisher Scientific, Inc., Multiskan plate reader and concentrations were determined via a standard curve.

3.2.4 Filter assisted sample preparation procedure (FASP) and tryptic digest

For the denaturation, reduction, alkylation and blocking of cysteine residues, the filter assisted sample preparation procedure (FASP) by Wisniewski *et al.* (2009) was used. A volume containing 300 µg proteins was transferred to a low-bind centrifuge tube (Protein loBind tube, Lasec), and 0.1 volumes of 50 mM Tris (2-carboxylethyl)-phosphine hydrochloride (TCEP) was added, followed by incubation in a heating block at 60 °C for 1 h to reduce cysteine disulphide bonds. The reduced protein sample was then transferred to a 30 kDa molecular weight cut off centrifugal Amnicon filter (Merck, USA) and inserted into the supplied collection tube, where the volume was reduced to 30 µl by centrifugation at 10,000 *x g* at room temperature. The sample was then incubated for 15 min at room temperature with 100 µl of 8 M urea in 0.5 M triethylammonium bicarbonate (TEAB), pH 8.5 containing 15 mM methyl methanethiosulfonate (MMTS) to block cysteine residues. To reduce the concentration of SDS, four washes with 8 M urea in 0.5 M TEAB was performed. In each wash the minimum volume of retentate was left in the filter before the next wash commenced. In a similar fashion, two washes with 0.5 M TEAB was then carried out to reduce the concentration of urea to an acceptable level (approximately 1 M).

For digestion of protein to peptide, proteomics-grade modified trypsin (Trypsin Gold, MS grade, Promega, USA) in 40 µl of 0.5M TEAB was added to samples at a trypsin: protein ratio of 1:100 (v/v). Optimal trypsin activity occurs at an alkaline pH, thus the pH was tested beforehand using pH strips, and adjusted to approximately pH 8-9 with 0.5 M TEAB if necessary. The tryptic digests were allowed to proceed overnight at 37 °C in a temperature incubator under sealed air-tight conditions to prevent evaporation.

3.2.5 iTRAQ labelling

Subsequent to incubation, protein tryptic digests were collected through centrifugation at 10,000 *x g* and transferred to new low-bind centrifuge tubes where each sample was concentrated down to 20 µl using a Savant SC110 Speed-Vac (Thermo Fisher Scientific, Inc.). For labelling of digested peptide, an 8-plex iTRAQ system was used (AbSciEx, USA). The iTRAQ tags used (113, 114, 115, 116, 117, 118, 119 and 121) for each respective treatment (hydrated and dehydrated conditions, including internal controls) is displayed in Figure 3.1. The labels were reconstituted with proteomics grade isopropanol and added to each sample, mixed by vortexing and left to incubate at room temperature for 2 h.

3.2.6 Peptide purification and OFFGEL fractionation

Once labelling had occurred, the contents for each labelled peptide sample were pooled together and reduced to approximately 30 µl by vacuum concentration using a speed-vac before being prepared for de-salting and purification on C-18 Spin Columns (Pierce, USA) according to the manufacturer's guide. The samples were reconstituted in 5% (v/v) acetonitrile (ACN) and 0.5% (v/v) Trifluoroacetic acid (TFA) and loaded onto columns pre-activated and equilibrated with 50% (v/v) ACN and 5% (v/v) ACN containing 0.5% (v/v) TFA, respectively, before centrifugation at 9,000 *x g* for 1 min at room temperature. Once peptides were bound to the column, contaminants were removed by washing twice with equilibration buffer (5% (v/v) ACN containing 0.5% (v/v) TFA), before eluting into clean low-bind centrifuge tubes using 70% ACN (v/v) with 0.1% (v/v) formic acid (FA). The purified peptide samples were then dried by vacuum concentration using a speed-vac before proceeding to OFFGEL fractionation.

For separation of labelled peptide samples according to their isoelectric points (p*I*), the 3100 OFFGEL fractionator (Agilent Technologies, USA) with a 12-well setup was used. The dried peptide samples were dissolved in a total volume of 1.8 ml 1.25X peptide OFFGEL rehydration solution (6% (v/v) glycerol, 1.25% (v/v) carrier ampholytes, at pH 3-10 (Sigma-Aldrich, Inc.)). Prior to OFFGEL fractionation, the frames for well-formation were assembled and two 13 cm immobilized pH gradient (IPG) strips (GE Healthcare, USA), with a linear pH 3-10 range, were left to rehydrate in 40 µl 1.25X peptide OFFGEL rehydration solution for 15 min according to the Agilent 3100 Quick Start Guide. Following IPG strip rehydration, 150 µl of re-solubilised peptide sample in 1.25X peptide OFFGEL rehydration solution was loaded in duplicate onto separate IPG strips into each of the 12 wells. Peptide electro-focusing was then performed using the pre-loaded OGPE12 program for peptide fractionation until a voltage of 20 $kV.h^{-1}$ was reached.

After electro-focusing, all the 12 peptide fractions belonging to each strip (2 in total) were retrieved and combined per fraction (fraction 1 in strip 1 pooled with fraction 1 in strip 2 and so on) into low-bind centrifuge tubes. The respective pooled fractions were purified using C-18 columns as described above to remove all traces of glycerol and contaminating substances. The digested, labelled, fractionated and purified peptide samples were then ready for analysis by ESI-Q-tof-MS/MS mass spectrometry.

3.2.7 Mass spectrometry settings

MS/MS analysis was carried out on each of the 12 purified peptide fractions using an Agilent 6530 quadrupole-time of flight (Q-TOF) mass spectrometer fitted with a Polaris HR 3 µm C18 high pressure liquid chromatography (HPLC)-Chip Cube source (Agilent Technologies, USA). The chip was equipped with a 75 µm x 150 mm analytical column and a 360 nl Zorbax enrichment column connected online to the 1200 Series nanoflow HPLC *via* an orthogonal spray HPLC-Chip/MS interface (Agilent Technologies, USA). Both systems were controlled by MassHunter Workstation Data Acquisition for Q-TOF (Agilent Technologies, USA).

Approximately 2 µg peptides were re-suspended in 1% (v/v) ACN and 0.1% (v/v) FA and loaded onto the trapping column at 1.6 µl. min^{-1} with the chip switched to enrichment and using the capillary pump. After loading, the chip was then switched to separation mode and peptides were eluted from the enrichment column and run through the separation column during a 1 h gradient (from 1% (v/v) ACN, 0.1% (v/v) FA to 90% (v/v) ACN, 0.1% FA) directly into the mass spectrometer. The mass spectrometer was run in positive ion mode, and MS scans were run over a range of *m/z* 200 to 1700 at a rate of seven spectra. sec^{-1}. MS/MS scans were run over a range of *m/z* 90 to 1700 at a scan rate of 2.50 spectra. sec^{-1} and a narrow (~1.3 amu) isolation width. Precursor ions were selected for auto MS/MS at an absolute threshold of 1000 and a relative threshold of 0.001, with a maximum of ten precursors per cycle, and active exclusion set at 1 spectrum and released after 1.5 min. Precursor charge-state selection and preference was set to 2+, 3+, and >3+, and precursors were sorted by abundance only.

3.2.8 Mass spectra data preparation

The raw mass spectra data files (.d format) retrieved from Agilent MassHunter software (Agilent Technologies, USA) were firstly converted to .mzML format followed by conversion to .mgf file formats, using the open source software, MSConvert available from the ProteoWizard (version 1.6.0) package (Kessner *et al.*, 2008).

The processed .mgf files were imported into PEAKS Studio (Bioinformatics Solutions Inc., version 6.0) developed by Ma *et al.* (2003) and the 'data refine' tool with default parameters (parent ion *m/z* tolerance at 0.1, retention time tolerance window of 30 sec, precursor charge correction, no merged scans and no filtering), were used to produce improved fragmentation, better signal-to-noise ratio and enhance reporter ion intensities.

3.2.9 Database selection

In order to have a comprehensive database search of tef proteins, two databases were selected to match proteins sequences to the iTRAQ generated mass spectra:

1.) The *Liliopsida* (all monocotyledonous plants) database available from UniProtKB Swiss-Prot/TREMBL (http://www.uniprot.org/downloads). The database was accessed on the 3 July 2013 in order to download all monocotyledonous plant protein sequences (reviewed sequences) in FASTA format.

2.) The Tef Extended transcriptome database converted to protein sequences (in FASTA format), available from the Tef Improvement Project (http://www.tef-research.org/genome.html), where the tef genome, transcriptome and proteome with annotations can be found. Access to the Tef Extended transcriptome database was kindly provided by the Tef research group at the University of Bern, Switzerland in August 2013, before being made available to the general public.

Both databases were verified in PEAKS Studio 6.0 and used in all subsequent searches.

3.2.10 Database searching

All generated iTRAQ mass spectra from tef protein samples were subjected to *de novo* sequencing and database searching with PEAKS Studio 6.0 software with the *Liliopsida* and Tef Extended databases. Two separate searches using PEAKS Studio 6.0 software with the *Liliopsida* and Tef Extended databases were performed. *De novo* sequencing was initiated with the following parameters (parent ion of 20.0 ppm using monoisotopic mass, a fragment ion of 0.1 Da and with enzymatic cleavage using trypsin). Database searching was employed with the following parameters (parent mass error tolerance of 20.0 ppm, fragment mass error tolerance of 0.1 Da, pre-cursor mass search type set as monoisotopic, selection of trypsin as enzyme used, non-specific cleavage set at 1, maximum missed cleavages per peptide set at 2, fixed modifications set at iTRAQ 8-plex (K, N-term) and beta-methylthiolation, variable modifications set at iTRAQ 8-plex (Y) and Oxidation (M) with max variable PTM per peptide set at 3). A concatenated decoy database for both the *Liliopsida* and Tef Extended databases were automatically generated by PEAKS Studio 6.0 when searches were implemented and further used to determine false discovery rates. The quantification results were then filtered so that the false discovery rate (FDR) was less than 1%, had a peptide and protein -10logP score of 20 or more and only considering proteins with 2 or more unique peptides. The resulting data were then auto-normalised using the PEAKS auto-normalisation tool to correct for channel bias and exported as .csv format to be used in further downstream analysis.

3.2.11 iTRAQ data processing – protein quantitation and statistical analysis

The output data generated from PEAKS was manually edited and refined before being analysed as shown in Figure 3.2.

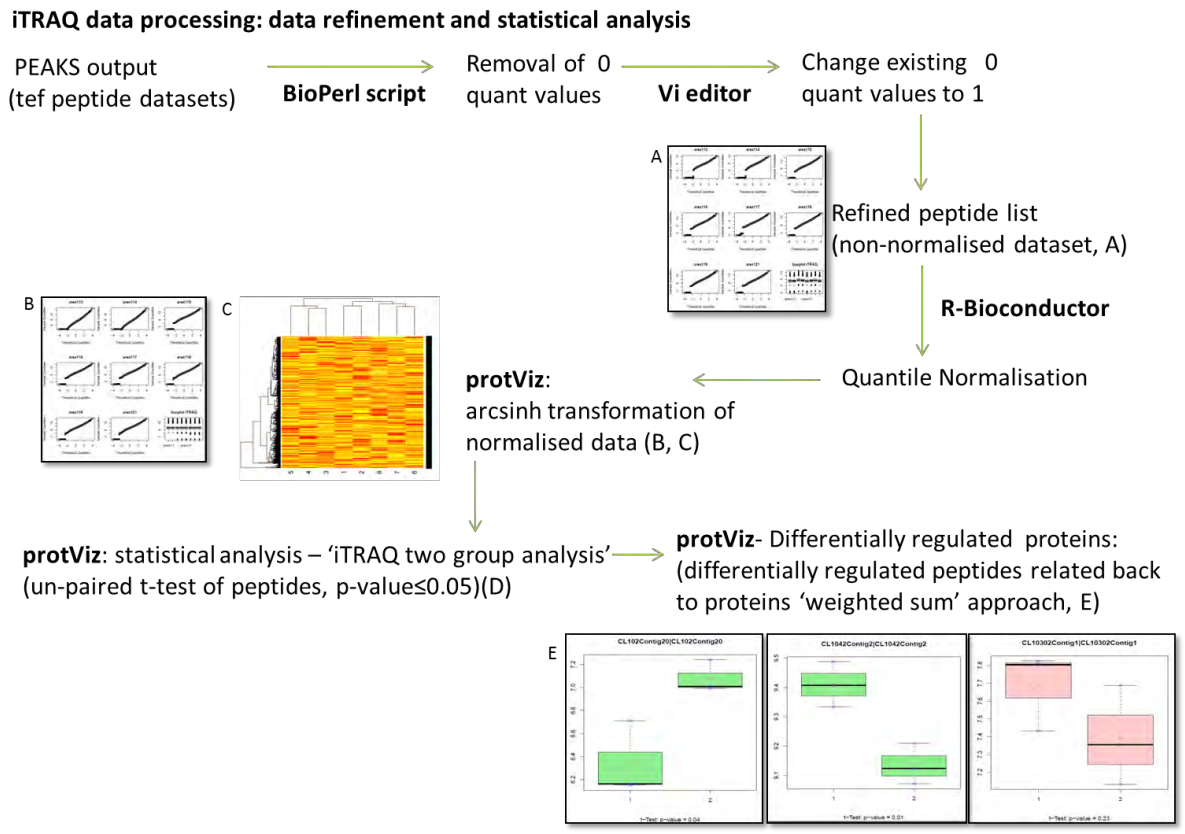

Fig. 3.2 Steps to tef iTRAQ data processing for peptide data refinement and statistical analysis, showing tools used to achieve each step. Data refinement and statistical analysis performed as follows: peptides with only 1 or 2 expression values present (of the 3 for each hydrated or dehydrated treatment) in the labelled channels (115 to 121) were removed. Peptides with a zero were kept if the zero values were found consecutively in the hydrated (115-117) or dehydrated (118-121) labelled channels (i.e. 115-117 had values of 0 while 118-121 had quantitative expression values or vice versa). These peptide expression values of zero were then changed to 1 to avoid later numerical errors (A), refined non-normalised peptide list subjected to quantile normalisation and inverse hyperbolic sine transformation (B, C), two-group comparison and statistical analysis of normalised, transformed peptide quantitative expression values (D), followed by conversion of respective peptides to corresponding proteins, displayed as box-plots (E).

3.2.11.1 Data refinement

Data refinement was conducted as follows: firstly, the peptide output list generated from PEAKS database searching was used, as opposed to the protein list, as processing and analysing peptide quantitative expression data is easier than working with whole proteins (Panse and Grossmann, 2012). The peptide list generated from PEAKS corresponds to the proteins identified from the Tef Extended and *Liliopsida* database searches with appropriate FDRs and threshold scores (-10logP scores) and editing of this list can be related back to the proteins identified.

Secondly, using a BioPerl script (Stajich *et al.*, 2002), peptides with only 1 or 2 expression values present (of the 3 for each treatment) in the labelled channels (115 to 121) were removed. Peptides with a zero value were kept if the zero values were found consecutively in the hydrated (115-117) or dehydrated (118-121) labelled channels, respectively. Values of zero in both the labelled channels 113 and 114, used as internal controls to observe technical variance between samples, were also retained. In order to avoid

problems with numerical computations in the downstream analysis, the remaining zeros were changed to 1 (excluding labelled channels 113 and 114). These changes resulted in a refined peptide list with non-normalised quantitative expression values (Fig. 3.2A).

For normalisation of peptide quantitative expression data, quantile normalisation was employed using the R-Bioconductor program (Gentleman *et al.*, 2004). In quantile normalisation, for each sample, intensities are rearranged from smallest to largest in columns and then averaged across rows. The averaged values then replace the original intensities, followed by rearrangement of averaged values in the original order (Chung *et al.*, 2014). This causes uniformity in the distribution of intensities across reporter ion channels and allows clear inferences to be made at a later stage (Fig. 3.2B).

3.2.11.2 protViz: for visualising and analysis of proteomic mass spectrometry data

The refined normalised peptide list was then transformed using an inverse hyperbolic sine function (arcsinh) (Fig. 3.2B, C) and statistically analysed using the R-Bioconductor program (Gentleman *et al.*, 2004) with the protViz package (http://cran.r-project.org/web/packages/protViz/index.html), developed by Panse and Grossmann (2012). Arcsinh transformation of normalised quantitative expression values is often preferable to Log_2 transformation because its use avoids errors that occur when using Log_2 transformation of zero values as well as the generation of negative values from the Log_2 transformation of values between 0 and 1.

Subsequent to normalisation and transformation steps, peptide quantitative expression data was laid out in a two group comparison manner using the 'iTRAQ two group analysis' function in protViz. In this comparison, the expression values from the labelled channels were placed in two groups, whereby group 1 consisted of all the hydrated labelled channels (115-117) and group 2 consisted of all the dehydrated labelled channels (118-121). Statistical testing was performed through an independent samples t-test (unpaired) that assigns a p-value to each individual peptide identity and tests for a significant difference (p-value ≤ 0.05) through the two-group comparison test between hydrated (control) and dehydrated (experimental) quantitative expression values (Fig. 3.2D). While the analysis was performed on each individual peptide identity, the output was given in such a manner that statistical significance for change in quantitative expression is observed in protein form. Thus, the peptides corresponding to the designated proteins are then stacked together through a 'weighted sum' approach to provide the overall change in quantitative expression between individual proteins where the result is displayed as box-plots for better interpretation (Fig. 3.2E).

3.2.12 Protein identification

All proteins matched to the Tef Extended and *Liliopsida* databases with PEAKS Studio 6.0 were annotated using Blast2GO version 2.8 (Conesa *et al.*, 2005). To provide protein descriptions, both datasets (in FASTA format) consisting of Tef Extended and *Liliopsida* matched proteins were searched against the UNIPROTKB/SwissProt database using the BLASTP algorithm with the following parameters: report a maximum of twenty blast hits, with a blast expect value of $1e^{-3}$ and minimum high scoring segment pairs (HSPs) length equal to 33. FASTA sequences for Tef Extended and *Liliopsida* matched proteins were retrieved from either database using a shell script (SH file for extracting database contents) written for extracting FASTA files.

3.2.13 OrthoMCL database search tool

OrthoMCL database searching was performed on the tef differentially regulated protein datasets, the Tef Extended (TE), Tef Extended-unique (TEU) and Monocot-unique (MU). The tool works by grouping proteins into "orthologous groups" based on sequence similarity through reciprocal BLAST and normalisation techniques, followed by the clustering of normalised BLAST scores using Markov clustering (Enright *et al.*, 2002; Fischer *et al.*, 2011). The search was performed with FASTA sequences of each of three differentially regulated datasets using default search parameters of: BLASTP e-value cut-off of $1e^{-5}$, maximum alignments of 50 and applying a low complexity filter.

3.2.14 Venn diagram generation

A Venn diagram for the tef differentially regulated protein datasets was generated using the software tool, Lucid Chart, to display protein groupings.

3.3 Results and Discussion

3.3.1 Preamble

The objective of this chapter was to employ the iTRAQ method of protein analysis to observe differential regulation of the tef proteome in response to controlled dehydration stress. Although these results are described in sections 3.3.2 and 3.3.4, a few noteworthy points need to be highlighted.

3.3.1.1 Manipulation and refinement of peptides instead of proteins

The starting point of the analysis was the list of peptides generated from PEAKS as opposed to the list of proteins generated through database searching. This allowed interrogation of a more complete dataset that was more representative of the proteomic profile under study. This is not an uncommon approach and has been used by many researchers in the field of mass spectrometry-based proteomics (explained in Choe *et al.*, 2007; Karp *et al.*, 2010; Panse and Grossmann, 2012; Thompson *et al.*, 2012). One of the potential concerns with working with a list of peptides instead of proteins is the challenge of protein inference (Cappadona *et al.*, 2012), where the generated list contains both unique and non-unique peptides matched against the chosen database for protein identification. In simpler terms, the list generated, contains peptides that are both unique to a certain protein (belonging to that protein only) and non-unique or shared (belonging to said protein and other proteins as well) when searching against the chosen database. This concern, however, is more than adequately addressed by using appropriate FDR thresholds, by employing stringent estimation of error rates, so that only valid peptide identities meeting the FDR threshold requirements are detected (Gupta and Pevzner, 2009; Karp *et al.*, 2010) and used for protein analysis. Furthermore, the analysis of both uniquely and non-uniquely scanned peptides would be more representative of the proteins changing in response to dehydration stress.

For this reason three complete datasets are displayed (in section 3.3), where in (i) proteins were matched to the Tef Extended database (from here on referred to as the TE dataset) containing both unique and non-uniquely matched peptides; (ii) proteins were matched to the Tef Extended database but only containing uniquely matched peptides (from here on referred to as the TEU dataset) and (iii) proteins matched to the *Liliopsida* (all monocots) database focusing on proteins made up of unique peptides only (from here on referred to as the MU dataset). The *Liliopsida* database is a vast (encompassing all monocotyledonous plants) and well-annotated database that resulted in large amounts of peptides being identified. For this reason and to avoid repetition and redundancy within the list, only uniquely scanned peptides were retained to allow the number of proteins identified, more manageable for data analysis.

3.3.2 iTRAQ data pre-processing and quality control

To examine whether data refinement tools were efficient in transforming PEAKS peptide output into an acceptable format for efficient statistical analysis, various quality checks were performed using the R-Bioconductor program (Gentleman *et al.*, 2004) and protViz package (Panse and Grossmann, 2012).

3.3.2.1 Data refinement and protViz

The data refinement steps were performed to reduce noise within the datasets and to observe accurate differential quantitative expression. Thus, the removal of zero values not consecutively belonging to the hydrated or dehydrated labelled channels (115-121), followed by the conversion of remaining zero values to 1 (except in the labelled channels 113 and 114) allowed the generation of a refined peptide quantitative expression dataset. These peptide quantitative expression values were then subjected to normalisation in the form of quantile normalisation to correct for sample bias as a result of sample preparation and mass spectrometry steps, followed by transformation procedures using archsinh function in protViz. The quality check results for these steps before statistical analysis are displayed in Figures 3.3 and 3.4.

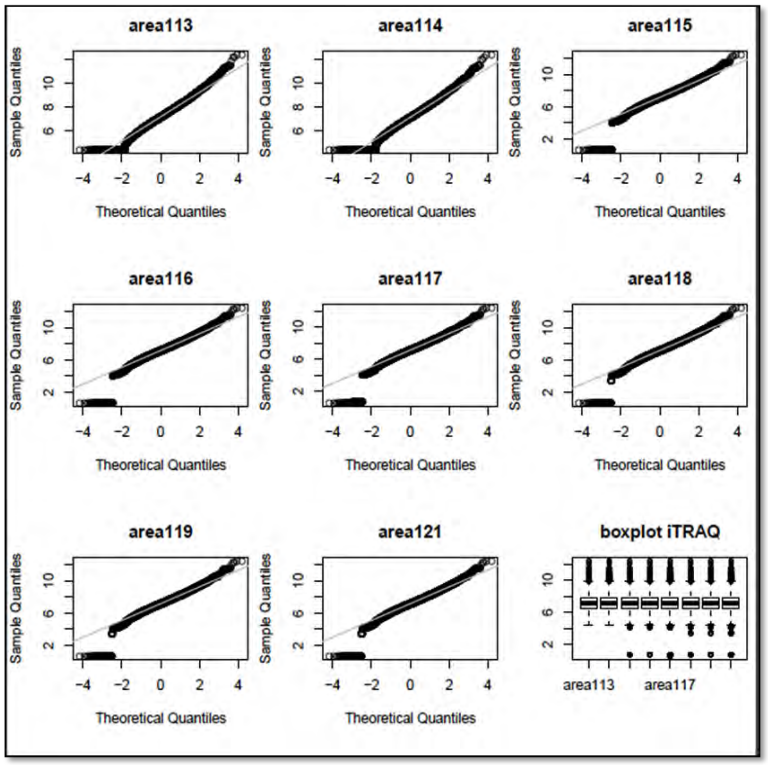

Fig. 3.3 Sanity Check (Q-Q plots) of all labels (113-121) in iTRAQ experiment that have been quantile normalised and archsinh (inverse hyperbolic sine) transformed to observe reporter ion channel (label) distributions. The last figure is a boxplot for all individual channels showing normalised channels.

To observe if reporter ion channels (labels 113-121) were normally distributed after data refinement steps, a sanity check (Q-Q plots) of theoretical quantiles plotted against the sample quantiles was conducted. In Figure 3.3, the theoretical quantiles (averaged values) and sample quantiles (actual

expression values) fall more or less on a straight line, except for zero-containing labelled channels 113 and 114 (internal controls). The quantile normalisation step was thus successful in making the distribution of intensities identical across samples (Chung *et al.*, 2014) and further facilitated medians across reporter ion channels to be approximately the same (9th figure in Fig. 3.3). In addition, a correlation test in the form of a heat map (cluster analysis) was performed to observe whether the labelled channels (113-121) clustered together after data refinement (Fig 3.4).

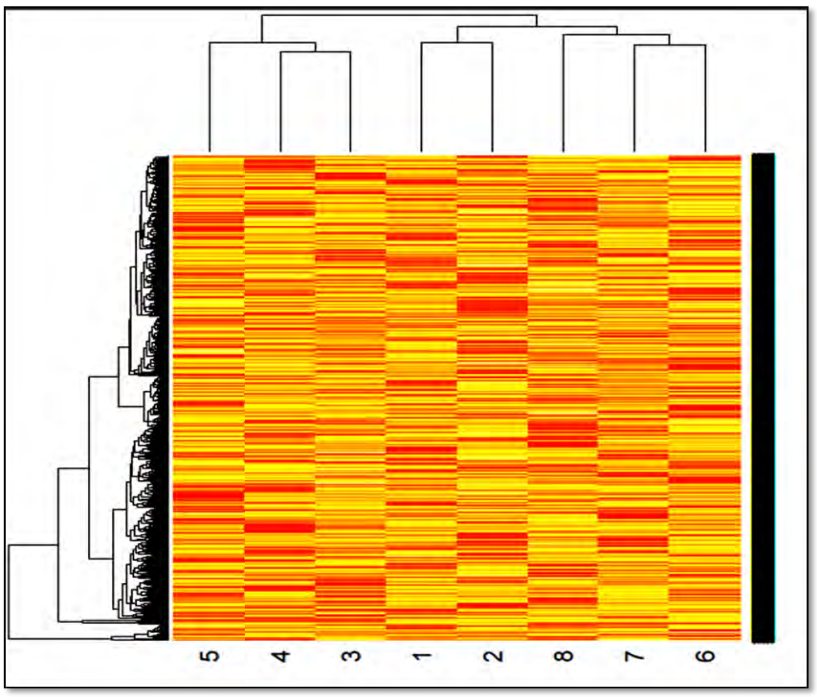

Fig. 3.4 Heat map (cluster analysis) of the labelled channels in iTRAQ experiment after data refinement steps. Channels 1 and 2 represent labels 113 and 114, respectively (pooled samples containing equal volumes of all hydrated and dehydrated proteins used in iTRAQ experiment). Channels 3, 4 and 5 represent the hydrated labelled peptides 115, 116 and 117, respectively, while channels 6, 7 and 8 represent dehydrated labelled peptides 118, 119 and 121, respectively.

As seen by the heat map (Fig. 3.4), the respective reporter ion channels are clustered according to experimental design. The internal control samples (labels 113 and 114 represented by channels 1 and 2, respectively) are clustered in the middle of the heat map, while the hydrated-control labels (115, 116 and 117 represented by channels 3, 4 and 5, respectively) cluster together and the dehydrated-experimental labels (118, 119 and 121 represented by channels 6, 7 and 8, respectively) cluster together (Fig. 3.4). Furthermore, the data refinement steps (quantile normalisation and archsinh transformation) of peptide quantitative expression values produced clustering consistent with the hydrated and dehydrated treatments.

3.3.2.2 protViz for statistical analysis

The protViz R-package (Panse and Grossmann, 2012), was employed for statistical analysis of quantitative expression data in order to observe proteins changing in response to dehydration stress. The

'iTRAQ two group analysis' function was used to observe the statistical significant difference in quantitative expression of values in group 1 (hydrated-control) to group 2 (dehydrated-experimental) of each peptide through an independent samples, unpaired t-test. Statistical significance was given in the form of a p-value (p-value ≤ 0.05) for differentially regulated proteins. In this approach, each peptide undergoes an independent samples t-test for statistical significance, followed by stacking each peptide together through a 'weighted sum' approach to give a final p-value for the corresponding protein. The output is given as a text file and box plots (Fig. 3.5) of proteins changing in differential protein expression in response to dehydration stress.

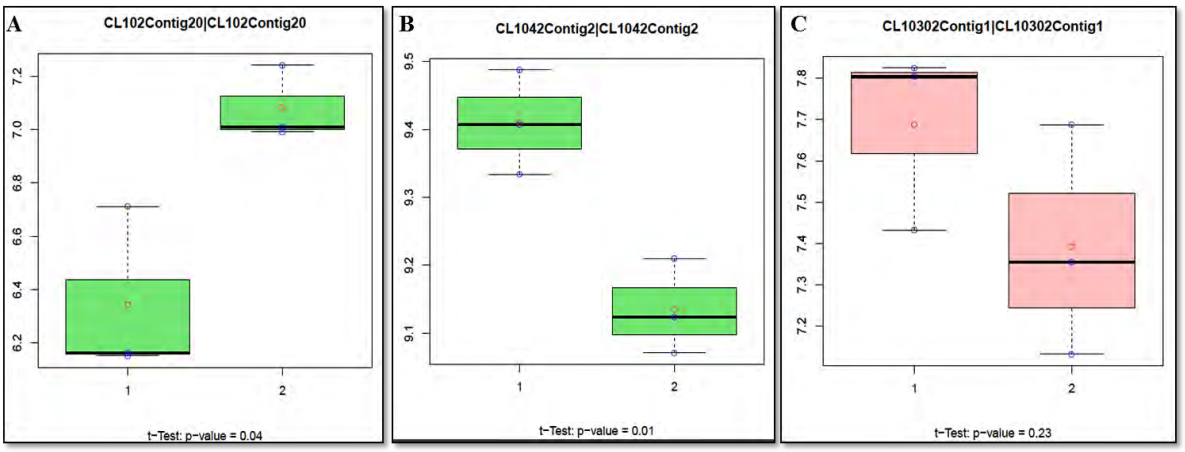

Fig. 3.5 Two-group comparisons (box-plots) between hydrated vs. dehydrated proteins. Statistically significant proteins are shown in green (p-value ≤ 0.05), by either being displayed as high abundance proteins (A) or low-abundance proteins (B), in response to dehydration stress. Proteins not statistically significantly changing in response to dehydration stress (p-value > 0.05), are shown in pink (C).

When multiple proteins are tested in the form of repetitive t-tests, the number of false-positive test results should be limited by multiple testing correction (Cappadona *et al.*, 2012). However, correction for multiple testing was not performed on our datasets, because stringent pruning was used to refine the data before analysis and the data would be biologically validated using western blots and physiological assays to ascertain increased or decreased abundance of proteins (discussed in Chapter 5).

3.3.3 Identification of differentially regulated proteins

Through the use of the protViz package and statistical testing, a total of 211 out of 5727 identified proteins for the TE dataset were found to be statistically significantly different in quantitative expression, where 97 proteins were found in high-abundance and 114 proteins were found in low-abundance in response to dehydration stress (Tables 3.1 and 3.2, respectively). For the TEU dataset a total of 111 out of 2656 identified proteins were statistically significant with 44 high-abundance proteins and 67 low-abundance proteins (Tables 3.3 and 3.4, respectively). While for the MU dataset, a total of 174 out of 4328 identified proteins were statistically significant with 85 high-abundance proteins and 89 low-abundance proteins (Supplementary Tables S3.1 and S3.2, respectively). The text files (excel spread

sheets) of all identified proteins, box-plots and quality control figures for all protein datasets (TE, TEU and MU) have been listed in the supplementary material section (section 3.5) and included as supplementary files (files S3 to S5) on the online cloud storage application, Dropbox. A link to access these supplementary data files is provided in section 3.5.

Table 3.1 TE high-abundance proteins in response to dehydration stress. Protein ID: protein identifier; protein description: identified protein; hydrated: averaged quantitative expression values belonging to hydrated labels 115-117 (a); dehydrated: averaged quantitative expression values belonging to dehydrated labels 118-121 (b); fold change: change in quantitative expression between hydrated and dehydrated values (b/a), where values > 1 display an increase in fold change; p-value: associated p-value for statistical significance (p-value ≤ 0.05).

Protein ID	Protein Description	Hydrated-a	Dehydrated-b	Fold change	p-value
CL1Contig10009	---NA---	7.22	8.09	1.12	0.001
CL5492Contig2	fructose-bisphosphate aldolase isozyme	10.70	10.86	1.02	0.001
CL856Contig3	40S ribosomal protein S28	9.17	9.33	1.02	0.001
CL856Contig4	40S ribosomal protein S28	9.17	9.33	1.02	0.001
CL68Contig25	peroxidase 3-rare cold-inducible protein	7.14	7.73	1.08	0.002
CL4104Contig2	gras family protein 2	0.69	4.68	6.75	0.003
CL3156Contig1	---NA---	0.69	6.05	8.73	0.004
CL6974Contig4	---NA---	0.69	6.05	8.73	0.004
CL715Contig2	protease, reverse transcriptase, endonuclease	9.03	9.43	1.04	0.004
ENO2_ERATE	---NA---	11.01	11.18	1.02	0.004
CL8690Contig2	---NA---	0.69	5.34	7.70	0.005
CL1942Contig1	---NA---	0.69	4.53	6.54	0.006
CL2699Contig6	monodehydroascorbate isoform 2	0.69	4.53	6.54	0.006
isotig02308	probable monodehydroascorbate isoform 2	0.69	4.53	6.54	0.006
CL2976Contig3	---NA---	0.69	4.94	7.12	0.007
CL68Contig6	peroxidase 3-rare cold-inducible protein	0.66	4.49	6.82	0.008
CL799Contig2	hua2-like protein 2	9.35	9.58	1.02	0.008
CL8983Contig3	---NA---	0.66	4.49	6.82	0.008
CL2991Contig2	hydroxyphenylpyruvate reductase	10.22	10.29	1.01	0.009
CL3629Contig1	poly polymerase i	0.69	5.13	7.41	0.013
CL3629Contig2	poly polymerase i	0.69	5.13	7.41	0.013
CL7746Contig3	---NA---	0.69	5.04	7.28	0.013
CL873Contig3	cyclin-p4-1	0.69	5.04	7.28	0.013
CL2289Contig1	---NA---	6.60	7.55	1.14	0.014
CL2761Contig4	red chlorophyll catabolite reductase	8.78	9.12	1.04	0.014
CL5492Contig1	fructose-bisphosphate aldolase isozyme	10.66	10.84	1.02	0.014
CL1Contig421	calcium-dependent protein kinase 5	7.83	8.40	1.07	0.015
CL1Contig7756	serine carboxypeptidase-like 51	9.51	9.77	1.03	0.015
CL3374Contig5	---NA---	0.69	5.41	7.80	0.016
CL57Contig23	---NA---	0.69	5.41	7.80	0.016
CL4404Contig1	---NA---	4.75	6.49	1.37	0.017
CL8759Contig1	proteasome subunit beta type-3	4.75	6.49	1.37	0.017

CL8759Contig2	proteasome subunit beta type-3	4.75	6.49	1.37	0.017
CL1Contig7029	protease, reverse transcriptase, endonuclease	6.94	7.76	1.12	0.018
CL1Contig4553	chlorophyll a-b binding protein	7.16	7.82	1.09	0.019
CL2228Contig1	s phase cyclin a-associated protein	9.04	9.19	1.02	0.019
CL124Contig7	---NA---	8.82	9.24	1.05	0.02
CL415Contig1	glutathione hydrolase 3	6.58	7.51	1.14	0.02
CL124Contig2	---NA---	8.30	8.72	1.05	0.021
CL3894Contig4	leucoanthocyanidin dioxygenase	0.69	5.49	7.92	0.021
CL1498Contig7	guanosine nucleotide diphosphate dissociation inhibitor 2	5.56	6.77	1.22	0.022
CL413Contig11	u-box domain-containing protein 4	7.18	7.96	1.11	0.022
ENO3_ERATE	---NA---	10.96	11.08	1.01	0.022
CL1888Contig1	glycerophosphodiester phosphodiesterase gdpdl3	11.40	11.61	1.02	0.023
CL1888Contig2	glycerophosphodiester phosphodiesterase gdpdl3	11.40	11.61	1.02	0.023
CL1498Contig4	guanosine nucleotide diphosphate dissociation inhibitor 1	6.53	7.06	1.08	0.024
CL1498Contig5	guanosine nucleotide diphosphate dissociation inhibitor 2	6.53	7.06	1.08	0.024
CL5028Contig3	plant intracellular ras-group-related lrr protein 6	8.98	9.21	1.03	0.024
isotig02787	guanosine nucleotide diphosphate dissociation inhibitor 2	6.53	7.06	1.08	0.024
CL3527Contig4	nucleolar complex protein 2 homolog	6.24	7.62	1.22	0.027
CL1Contig6763	---NA---	5.57	6.63	1.19	0.028
CL1224Contig6	gtp-binding protein sar1a	7.63	8.20	1.07	0.029
CL90Contig16	---NA---	7.59	8.02	1.06	0.029
CL18849Contig1	---NA---	5.79	6.54	1.13	0.031
CL4737Contig2	acetohydroxy-acid reductoisomerase	9.65	9.89	1.03	0.032
CL577Contig14	ubiquinol oxidase	9.39	9.56	1.02	0.032
CL24657Contig1	fructokinase-1	7.48	7.92	1.06	0.036
CL546Contig2	f-box protein skip24	5.26	6.16	1.17	0.036
CL7996Contig1	fructokinase-1	7.48	7.92	1.06	0.036
CL1073Contig1	peptidyl-prolyl cis-trans isomerase d	7.15	7.79	1.09	0.037
CL3347Contig4	delta-aminolevulinic acid	8.68	9.00	1.04	0.037
CL5Contig21	probable wrky transcription factor 19	7.72	8.19	1.06	0.037
CL7405Contig3	chlorophyll a-b binding protein cp24	10.71	10.85	1.01	0.038
CL136Contig17	---NA---	7.58	8.24	1.09	0.04
CL1Contig5054	f-box only protein 8	7.30	7.68	1.05	0.04
CL326Contig6	---NA---	4.92	6.56	1.33	0.04
CL3687Contig5	---NA---	4.92	6.56	1.33	0.04
CL4000Contig1	monodehydroascorbate reductase	9.99	10.26	1.03	0.04
CL4000Contig2	monodehydroascorbate reductase	9.99	10.26	1.03	0.04
CL4000Contig3	monodehydroascorbate reductase	9.99	10.26	1.03	0.04
CL4207Contig1	---NA---	4.92	6.56	1.33	0.04
CL445Contig4	elongator complex protein 6	4.92	6.56	1.33	0.04

CL445Contig6	elongator complex protein 7	4.92	6.56	1.33	0.04
CL4771Contig3	---NA---	4.92	6.56	1.33	0.04
CL4956Contig4	premnaspirodiene oxygenase	4.92	6.56	1.33	0.04
CL522Contig8	---NA---	4.92	6.56	1.33	0.04
CL6050Contig1	---NA---	4.92	6.56	1.33	0.04
CL7668Contig1	nadh dehydrogenase complex assembly factor 6	6.77	7.66	1.13	0.04
CL7668Contig2	nadh dehydrogenase complex assembly factor 6	6.77	7.66	1.13	0.04
CL837Contig7	cell division cycle protein 48 homolog	10.78	10.96	1.02	0.04
CL102Contig20	---NA---	6.34	7.08	1.12	0.041
CL5577Contig3	---NA---	7.34	7.84	1.07	0.041
CL1Contig7889	---NA---	8.42	9.03	1.07	0.042
CL2761Contig3	red chlorophyll catabolite reductase	8.59	8.97	1.04	0.042
CL836Contig11	probable polyamine transporter	7.78	8.38	1.08	0.042
CL4591Contig2	phosphatidylinositol n-acetylglucosaminyltransferase subunit a	8.95	9.18	1.03	0.043
CL61Contig20	probable ufm1-specific protease	8.15	8.64	1.06	0.043
CL61Contig7	probable ufm1-specific protease	8.15	8.64	1.06	0.043
CL4852Contig2	---NA---	9.18	9.68	1.05	0.046
CL2637Contig1	peroxisome biogenesis protein 6 (PEX6)	9.85	10.06	1.02	0.047
CL680Contig10	elongation factor tu gtp-binding domain-containing protein 2	7.43	8.05	1.08	0.048
CL680Contig5	elongation factor tu gtp-binding domain-containing protein 2	7.43	8.05	1.08	0.048
CL7065Contig1	---NA---	5.58	6.49	1.16	0.048
CL7065Contig2	---NA---	5.58	6.49	1.16	0.048
CL140Contig10	npk1-activating kinesin-1	8.98	9.27	1.03	0.049
CL4289Contig6	---NA---	8.29	8.58	1.03	0.049
isotig08284	protein disulfide isomerase-like 1	10.76	10.98	1.02	0.049

Table 3.2 TE low-abundance proteins in response to dehydration stress. Protein ID: protein identifier; protein description: identified protein; hydrated: averaged quantitative expression values belonging to hydrated labels 115-117 (a); dehydrated: averaged quantitative expression values belonging to dehydrated labels 118-121 (b); fold change: change in quantitative expression between hydrated and dehydrated values (b/a), where values < 1 display a decrease in fold change; p-value: associated p-value for statistical significance (p-value ≤ 0.05).

Protein ID	Protein Description	Hydrated-a	Dehydrated-b	Fold change	p-value
CL5604Contig1	2-methyl-6-phytyl-hydroquinone methyltransferase	9.36	8.94	0.96	0.001
CL977Contig4	---NA---	7.70	7.08	0.92	0.001
CL2349Contig3	protein dek	9.77	9.35	0.96	0.002
CL36Contig35	nad-dependent malic enzyme 59 kda	10.52	10.36	0.99	0.002
CL5457Contig2	---NA---	9.09	8.91	0.98	0.002
CL11972Contig1	metal tolerance protein 5	6.23	1.19	0.19	0.004
CL700Contig3	nad-dependent malic enzyme 62 kda	9.37	9.12	0.97	0.005
CL236Contig5	probable sucrose-phosphate synthase 2	9.46	9.21	0.97	0.006
CL236Contig6	probable sucrose-phosphate synthase 2	9.44	9.18	0.97	0.006
CL1456Contig11	s-adenosylmethionine decarboxylase proenzyme	5.33	0.69	0.13	0.007
CL1456Contig8	s-adenosylmethionine decarboxylase proenzyme	5.33	0.69	0.13	0.007
CL2948Contig2	haloalkane dehalogenase	9.85	9.63	0.98	0.007
isotig10649	s-adenosylmethionine decarboxylase proenzyme	5.33	0.69	0.13	0.007
CL13Contig40	rhodanese-like domain-containing protein 10	8.66	7.80	0.90	0.009
CL1595Contig2	---NA---	4.99	0.69	0.14	0.009
CL1042Contig2	alpha-glucan water	9.41	9.13	0.97	0.01
CL19Contig25	---NA---	8.92	8.29	0.93	0.01
CL7534Contig1	cellulose synthase-like protein a9	7.36	6.55	0.89	0.01
CL7716Contig2	---NA---	7.78	7.25	0.93	0.013
CL7716Contig3	Putative uncharacterized protein CysX	7.78	7.25	0.93	0.013
CL7582Contig1	ribosomal rna processing protein 36 homolog	8.47	7.75	0.92	0.014
CL14686Contig1	alliin lyase 1	5.51	0.69	0.13	0.015
CL2382Contig6	chlorophyll a-b binding	9.39	8.99	0.96	0.015
Locus_49_75_82	---NA---	11.12	10.92	0.98	0.015
CL1759Contig3	---NA---	7.61	7.19	0.94	0.016
CL1Contig8969	---NA---	10.54	10.40	0.99	0.016
CL14672Contig1	---NA---	8.33	7.87	0.94	0.017
CL16131Contig1	---NA---	8.33	7.87	0.94	0.017
CL456Contig16	---NA---	10.12	9.95	0.98	0.017
CL7059Contig2	cytochrome b561 domain-containing protein	8.00	7.36	0.92	0.017
CL9348Contig2	ras-related protein raba5c	9.07	8.73	0.96	0.017
CL1630Contig1	nadh azoreductase	9.33	9.12	0.98	0.018
CL3227Contig1	---NA---	5.69	0.69	0.12	0.018
CL3227Contig2	---NA---	5.69	0.69	0.12	0.018
CL58Contig14	---NA---	5.69	0.69	0.12	0.018

CL58Contig2	---NA---	5.69	0.69	0.12	0.018
CL467Contig12	---NA---	9.78	9.57	0.98	0.02
isotig23406	photosystem ii protein d1	10.59	10.33	0.98	0.02
CL1Contig3562	---NA---	8.20	7.53	0.92	0.022
CL3294Contig3	nadh-quinone oxidoreductase subunit	9.37	9.17	0.98	0.022
CL3294Contig4	nadh-quinone oxidoreductase subunit	9.37	9.17	0.98	0.022
CL3294Contig5	nadh-quinone oxidoreductase subunit	9.37	9.17	0.98	0.022
CL3294Contig6	nadh-quinone oxidoreductase subunit	9.37	9.17	0.98	0.022
CL6495Contig2	polyamine oxidase	8.91	8.42	0.95	0.022
comp294_c0_seq1	nadh-quinone oxidoreductase subunit	9.37	9.17	0.98	0.022
CL5963Contig1	60S ribosomal protein l5-1	10.95	10.86	0.99	0.023
CL8805Contig2	---NA---	9.60	9.14	0.95	0.023
CL5672Contig2	---NA---	8.49	8.03	0.95	0.024
CL6932Contig1	---NA---	9.87	9.60	0.97	0.024
Locus_954_4_4	---NA---	9.87	9.60	0.97	0.024
SYN3_ERATE	---NA---	9.96	9.63	0.97	0.024
SYN8_ERATE	---NA---	9.96	9.63	0.97	0.024
CL1805Contig10	protein dj-1 homolog b	8.06	7.57	0.94	0.025
CL1805Contig2	protein dj-1 homolog b	8.06	7.57	0.94	0.025
CL1Contig3266	---NA---	7.90	7.43	0.94	0.025
CL327Contig3	cbs domain-containing protein	7.85	7.37	0.94	0.025
Locus_2288_7_9	---NA---	9.38	9.05	0.97	0.025
CL2320Contig2	---NA---	7.27	6.85	0.94	0.026
CL977Contig1	---NA---	7.27	6.85	0.94	0.026
CL10226Contig1	---NA---	8.21	7.20	0.88	0.027
CL2336Contig7	---NA---	9.33	9.15	0.98	0.027
CL1Contig5286	metal tolerance protein 5	5.00	0.69	0.14	0.029
CL1Contig5699	metal tolerance protein 6	5.00	0.69	0.14	0.029
CL1Contig8303	metal tolerance protein 5	5.00	0.69	0.14	0.029
CL73Contig10	clathrin heavy chain 1	11.05	10.94	0.99	0.03
CL3204Contig2	---NA---	6.50	5.77	0.89	0.031
CL10162Contig3	---NA---	7.16	6.23	0.87	0.032
CL349Contig4	homeobox-leucine zipper protein roc6	7.24	6.70	0.93	0.033
CL349Contig7	homeobox-leucine zipper protein roc6	7.24	6.70	0.93	0.033
CL5826Contig1	long chain acyl- synthetase 4	10.35	10.18	0.98	0.033
CL5826Contig2	long chain acyl- synthetase 4	10.35	10.18	0.98	0.033
CL5942Contig6	---NA---	9.19	8.87	0.97	0.033
CL6511Contig2	v-type proton atpase subunit g1	9.82	9.58	0.98	0.033
CL7Contig43	---NA---	7.48	7.28	0.97	0.033
comp13984_c0_seq1	---NA---	8.33	7.85	0.94	0.035
CL1Contig3395	endoglucanase 7	8.19	7.50	0.92	0.036
CL1Contig3396	endoglucanase 7	8.19	7.50	0.92	0.036
CL1Contig3397	endoglucanase 7	8.19	7.50	0.92	0.036
CL3496Contig11	chlorophyll a-b binding protein	10.21	9.97	0.98	0.036
CL154Contig2	ankyrin repeat domain-containing protein	9.95	9.70	0.97	0.038

CL1Contig4279	---NA---	9.23	8.89	0.96	0.038
CL4622Contig2	rubredoxin	11.01	10.88	0.99	0.038
CL1Contig4299	histone-lysine n-methyltransferase setd3	8.28	7.84	0.95	0.039
CL1Contig4635	protease do-like 14	6.90	5.87	0.85	0.039
CL8953Contig2	---NA---	9.05	8.80	0.97	0.039
CL5563Contig3	---NA---	9.08	8.84	0.97	0.04
CL1Contig242	bax inhibitor 1	9.79	9.67	0.99	0.041
CL2736Contig1	ubiquinol-cytochrome-c reductase subunit ii	9.80	9.66	0.99	0.041
CL3528Contig3	golgin candidate 4	9.24	8.95	0.97	0.041
CL5380Contig1	glutathione s-transferase t3	8.68	8.50	0.98	0.041
CL94Contig6	---NA---	9.49	9.33	0.98	0.041
CL5774Contig2	r60s acidic ribosomal protein p0	9.92	9.78	0.99	0.042
CL3496Contig14	chlorophyll a-b binding protein 1b	10.11	9.95	0.98	0.043
CL3496Contig15	chlorophyll a-b binding protein 1b	10.11	9.95	0.98	0.043
CL811Contig3	myb-like transcription factor 1	7.47	6.96	0.93	0.043
CL3168Contig2	---NA---	8.47	8.06	0.95	0.044
Locus_393_4_9	---NA---	10.96	10.81	0.99	0.044
CL1Contig492	---NA---	9.24	8.99	0.97	0.045
CL1Contig6871	chlorophyll a-b binding protein	11.19	10.99	0.98	0.045
CL1Contig7112	hexose carrier protein hex6	8.75	8.57	0.98	0.045
CL185Contig19	probable disease resistance protein rf45	9.19	8.93	0.97	0.046
CL885Contig1	formin-like protein 3	9.72	9.54	0.98	0.046
Locus_61_5_6	---NA---	10.11	9.92	0.98	0.046
CL3496Contig13	chlorophyll a-b binding protein 1b	10.10	9.88	0.98	0.048
CL3496Contig17	chlorophyll a-b binding protein 1b	10.10	9.88	0.98	0.048
CL321Contig12	---NA---	8.27	7.66	0.93	0.049
CL785Contig5	---NA---	7.16	6.68	0.93	0.049
CL94Contig5	choline	9.49	9.35	0.99	0.049
CL131Contig9	---NA---	7.77	7.25	0.93	0.05
CL19309Contig1	---NA---	8.53	8.27	0.97	0.05
CL236Contig2	---NA---	8.87	8.56	0.97	0.05
CL236Contig9	---NA---	8.87	8.56	0.97	0.05
CL305Contig27	---NA---	8.53	8.27	0.97	0.05
CL7612Contig2	---NA---	8.53	8.27	0.97	0.05

Table 3.3 TEU high-abundance proteins in response to dehydration stress. Protein ID: protein identifier; protein description: identified protein; hydrated: averaged quantitative expression values belonging to hydrated labels 115-117 (a); dehydrated: averaged quantitative expression values belonging to dehydrated labels 118-121 (b); fold change: change in quantitative expression between hydrated and dehydrated values (b/a), where values > 1 display an increase in fold change; p-value: associated p-value for statistical significance (p-value ≤ 0.05).

Protein ID	Protein description	Hydrated-a	Dehydrated-b	Fold change	p-value
CL856Contig3	40S ribosomal protein S28	9.17	9.33	1.02	0.001
CL1127Contig6	morc family cw-type zinc finger protein 4	6.62	7.33	1.11	0.002
CL4104Contig2	gras family protein 2	0.88	4.68	5.32	0.002
CL1Contig10662	beta-glucosidase 10	5.90	6.63	1.12	0.004
CL14878Contig1	---NA---	6.19	6.81	1.10	0.005
CL5492Contig2	fructose-bisphosphate aldolase isozyme	10.69	10.87	1.02	0.005
CL714Contig6	gdp-mannose-epimerase 1	4.93	6.35	1.29	0.005
CL124Contig7	---NA---	8.03	8.57	1.07	0.006
CL6Contig61	---NA---	0.88	5.10	5.80	0.006
CL1Contig8614	magnesium-protoporphyrin ix monomethyl ester	7.93	8.30	1.05	0.007
CL2976Contig3	---NA---	0.88	4.95	5.63	0.007
CL8163Contig2	actin-1	6.48	7.47	1.15	0.007
CL243Contig5	serine threonine-protein kinase 11-interacting protein	8.07	8.72	1.08	0.008
CL5983Contig4	udp-glucose:cinnamate glucosyltransferase	0.88	4.76	5.41	0.01
CL7060Contig1	eukaryotic translation initiation factor 3 subunit h	7.78	8.01	1.03	0.011
CL1Contig7473	protoporphyrinogen	0.88	5.93	6.74	0.014
CL394Contig12	---NA---	7.82	8.39	1.07	0.014
CL1322Contig3	ubiquinol oxidase	6.26	6.92	1.11	0.015
CL1Contig4551	---NA---	0.88	5.40	6.14	0.016
CL1Contig4553	chlorophyll a-b binding protein	6.71	7.42	1.11	0.017
CL7068Contig1	---NA---	0.88	5.54	6.29	0.017
CL8759Contig1	proteasome subunit beta type-3	4.72	6.50	1.38	0.018
CL1533Contig11	uncharacterized oxidoreductase	7.33	7.77	1.06	0.019
CL3894Contig4	leucoanthocyanidin dioxygenase	0.88	5.50	6.25	0.021
SPEE4_ERATE	spermine synthase	5.28	6.38	1.21	0.021
CL1317Contig7	glyceraldehyde-3-phosphate dehydrogenase	0.88	4.92	5.59	0.023
CL136Contig17	---NA---	5.03	6.62	1.32	0.023
CL1498Contig7	guanosine nucleotide diphosphate dissociation inhibitor 2	5.51	6.78	1.23	0.024
CL5895Contig4	calcium sensing protein	7.72	7.95	1.03	0.024
CL4852Contig2	---NA---	8.90	9.52	1.07	0.025
CL4000Contig1	monodehydroascorbate reductase	9.93	10.23	1.03	0.027
CL61Contig20	probable ufm1-specific protease	7.49	8.09	1.08	0.028
CL13655Contig1	---NA---	5.12	6.54	1.28	0.03
CL1Contig3898	photosystem ii 10 kda protein	7.54	8.13	1.08	0.03
CL1Contig7889	---NA---	8.41	9.03	1.07	0.031

CL433Contig3	atpase family aaa domain-thorase	8.39	8.55	1.02	0.033
CL1030Contig10	nucleolar gtp-binding protein nsn1	5.79	7.17	1.24	0.037
CL102Contig20	---NA---	6.33	7.08	1.12	0.038
CL2Contig153	splicing factor u2af large subunit a	6.14	6.98	1.14	0.038
CL9427Contig2	---NA---	8.88	9.07	1.02	0.039
CL1371Contig3	nadph--cytochrome p450 reductase	5.76	6.58	1.14	0.041
CL1Contig167	extracellular lipase	8.56	8.92	1.04	0.042
CL1Contig10732	chlorophyll a-b binding protein	9.10	9.22	1.01	0.045
Locus_11848_5_6	lim domain-containing protein wlim1	7.21	7.68	1.07	0.05

Table 3.4 TEU low-abundance proteins in response to dehydration stress. Protein ID: protein identifier; protein description: identified protein; hydrated: averaged quantitative expression values belonging to hydrated labels 115-117 (a); dehydrated: averaged quantitative expression values belonging to dehydrated labels 118-121 (b); fold change: change in quantitative expression between hydrated and dehydrated values (b/a), where values < 1 display a decrease in fold change; p-value: associated p-value for statistical significance (p-value ≤ 0.05).

Protein ID	Protein description	Hydrated-a	Dehydrated-b	Fold change	p-value
CL111Contig6	---NA---	6.04	0.88	0.15	0.001
CL2546Contig3	40S ribosomal protein S2-4	7.82	7.36	0.94	0.001
CL5604Contig1	2-methyl-6-phytyl-hydroquinone methyltransferase	9.37	8.93	0.95	0.001
CL8805Contig2	---NA---	8.94	7.98	0.89	0.001
CL1042Contig2	alpha-glucan water chloroplastic precursor	9.22	8.76	0.95	0.002
CL2878Contig1	xanthine dehydrogenase	5.42	0.88	0.16	0.002
CL6845Contig2	---NA---	5.34	0.88	0.16	0.002
CL1Contig3694	---NA---	5.52	0.88	0.16	0.005
CL323Contig2	187-kda microtubule-associated protein air9	8.65	8.23	0.95	0.006
CL2349Contig3	protein dek	9.76	9.29	0.95	0.007
CL682Contig1	nuclease harbi1	7.22	5.73	0.79	0.01
CL8133Contig1	electron transfer flavoprotein subunit	7.44	6.18	0.83	0.01
CL1013Contig3	probable plastid-lipid-associated protein	7.27	6.05	0.83	0.011
CL1Contig962	---NA---	6.14	0.88	0.14	0.012
SECA9_ERATE	protein translocase subunit	9.92	9.68	0.98	0.013
CL16131Contig1	---NA---	8.34	7.87	0.94	0.014
CL3513Contig2	btb poz and math domain-containing protein 3	8.26	7.9	0.96	0.014
Locus_2288_7_9	---NA---	9.04	8.6	0.95	0.014
CL6495Contig2	polyamine oxidase	7.43	6.7	0.90	0.015
CL1670Contig2	intron-binding protein aquarius	7.4	6.28	0.85	0.016
CL1Contig8969	---NA---	10.54	10.33	0.98	0.016
CL7215Contig2	---NA---	8.94	8.42	0.94	0.016
CL14686Contig1	alliin lyase 1	5.4	0.88	0.16	0.017
CL14873Contig1	---NA---	4.94	0.88	0.18	0.017
CL1759Contig3	photosynthetic nadh subunit of subcomplex b	5.78	0.88	0.15	0.017
CL7716Contig2	---NA---	7.74	7.25	0.94	0.018
CL2491Contig2	---NA---	4.94	0.88	0.18	0.019

CL2706Contig2	probable gtp-binding protein	7.8	7	0.90	0.019
CL4711Contig1	e3 ubiquitin-protein ligase ring1a	8.01	7.61	0.95	0.019
CL1Contig8961	ubiquitin-nedd8-like protein rub1	8.18	7.74	0.95	0.021
CL316Contig3	---NA---	5.7	0.88	0.15	0.022
CL3294Contig6	nadh-quinone oxidoreductase subunit	9.37	9.17	0.98	0.024
CL1Contig3562	---NA---	8.14	7.54	0.93	0.026
CL227Contig12	beta-amylase	7.19	6.6	0.92	0.026
CL154Contig18	ankyrin repeat domain-containing protein	9.84	9.61	0.98	0.027
CL1662Contig3	---NA---	7.44	6.7	0.90	0.027
CL1805Contig7	protein dj-1 homolog b	8.1	7.57	0.93	0.027
CL338Contig6	---NA---	8.44	8.16	0.97	0.03
CL612Contig6	inactive ubiquitin carboxyl-terminal hydrolase 54	7.16	6.77	0.95	0.03
CL714Contig5	gdp-mannose -epimerase 1	7.2	6.93	0.96	0.032
CL7534Contig1	probable mannan synthase 9	7.17	6.54	0.91	0.033
CL3204Contig2	---NA---	6.44	5.77	0.90	0.034
CL909Contig5	phytoene chloroplastic flags	7	6.25	0.89	0.034
CL44Contig17	ras-related protein ric2	8.08	7.91	0.98	0.035
CL5788Contig2	---NA---	8.64	8.56	0.99	0.036
CL10469Contig1	nadp-dependent d-sorbitol-6-phosphate dehydrogenase	6.5	5.54	0.85	0.038
CL811Contig3	protein phr1-like 1	7.45	6.96	0.93	0.038
CL1Contig4635	protease do-like 14	6.8	5.87	0.86	0.039
CL2948Contig2	haloalkane dehalogenase	7.98	7.56	0.95	0.039
CL52Contig6	---NA---	7.03	6.39	0.91	0.039
CL1035Contig6	phosphoglycerate	9.75	9.57	0.98	0.04
CL5563Contig3	---NA---	9.04	8.76	0.97	0.04
Locus_49_80_82	chlorophyll a-b binding protein of lhcii type	10.48	10.36	0.99	0.04
CL1Contig1301	activating signal co-integrator 1 complex subunit 3	8.45	8.01	0.95	0.041
CL7698Contig2	50S ribosomal protein	9.53	9.172	0.96	0.041
CL1Contig4227	nadh-quinone oxidoreductase subunit	8.44	7.94	0.94	0.042
CL424Contig8	---NA---	6.74	5.89	0.87	0.043
Locus_471_2_7	glutamine	10.33	10.1	0.98	0.043
CL321Contig12	---NA---	8.24	7.66	0.93	0.044
CL10226Contig1	---NA---	8.04	7.08	0.88	0.045
CL36Contig12	nad-dependent malic enzyme 59 kda	8.87	8.57	0.97	0.046
CL3496Contig11	chlorophyll a-b binding protein 1b	10.14	9.84	0.97	0.047
CL9878Contig2	---NA---	8.54	8.27	0.97	0.047
CL113Contig3	40S ribosomal protein S20-2	7.87	7.27	0.92	0.048
CL1Contig2108	---NA---	9.07	8.32	0.92	0.048
CL3757Contig1	ankyrin repeat-containing protein	8.2	7.54	0.92	0.048
CL2414Contig3	---NA---	6.24	4.58	0.73	0.05

From the tables, the proteins identified from the Tef Extended database (TE and TEU) have usable quantitative information and are shown to be differentially regulated (p-values ≤ 0.05). However, a large proportion of the proteins present do not have protein descriptions. Approximately 67 and 63% of proteins were annotated and identified with Blast2GO tools (Conesa *et al.*, 2005; Gotz *et al.*, 2008) for TE high and low-abundance proteins, respectively (Tables 3.1 and 3.2) and 72 and 63% of proteins had descriptions for TEU high and low-abundance proteins, respectively (Tables 3.3 and 3.4). Furthermore, a number of proteins with the same protein descriptions and quantification values are repeated within the TE dataset (Tables 3.1 and 3.2). These proteins have arisen through alternative splicing and are spliced variants of the same protein, indicated by the same protein identifier but different suffixes (e.g. CL1Contig3395, CL1Contig3396, and CL1Contig3397 for endoglucanase 7 in Table 3.2). A total of 57 out of the 211 proteins (27%) found to be differentially regulated within the TE dataset (Tables 3.1 and 3.2), were spliced variants arising from the alternative splicing of 25 potential splice events (genes).

During this regulatory mechanism, primary transcripts or precursor-mRNAs with introns undergo alternative splicing to produce multiple transcripts from a single gene within the genome by using differential splice sites (Kazan, 2003). In this regard, the functional complexity of the transcriptome and diversity of the proteome are increased between plant cells and tissues (Kazan, 2003; Reddy, 2007), particularly during plant development and in response to environmental stimuli, such as biotic and abiotic stress conditions (Duque, 2011; Staiger, 2015). The different versions of mRNA transcripts generated through alternative splicing are later translated into different forms of a singular protein (protein isoforms etc.) that result in multiple entries of proteins with the same protein descriptions and quantification values, but different protein identifiers being detected by MS/MS. In the TEU differentially regulated datasets (Tables 3.3 and 3.4) and MU differentially regulated datasets (Table S3.1 and S3.2), however, no occurrences of spliced variants were present, presumably because only uniquely-matched peptides were used for protein identification, resulting in only one definitive protein entity per entry.

In addition, a suitably large number of proteins were found to be commonly identified in both the TE and TEU datasets (Tables 3.1 to 3.4). To name a few: 40S ribosomal protein S28, fructose-bisphosphate aldolase, monodehydroascorbate reductase, gras family protein 2 and leucoanthocyanidin dioxygenase found in high-abundance proteins (Tables 3.1 and 3.3) and 2-methyl-6-phytyl-hydroquinone methyltransferase, chlorophyll a-b binding protein, alliin lyase 1, protein dek and polyamine oxidase found in low-abundance proteins (Tables 3.2 and 3.4). Furthermore, the proteins leucoanthocyanidin dioxygenase and gras family protein 2 are among the proteins of high-abundance with descriptions that show the largest increase in fold change in quantitative expression (7.92 and 6.75, respectively, in Table 3.1; 6.25 and 5.32, respectively, in Table 3.3), while the protein, alliin lyase 1, shows the largest decrease in fold change in quantitative expression in low-abundance proteins (0.13 and 0.16, respectively, in Tables 3.2 and 3.4).

Many factors play a role in the attainment of differentially regulated proteins with useful qualitative information (Karp *et al.*, 2010), one of them being the extent of annotations and curations made within the chosen database (Carpentier *et al.*, 2008a; 2008b). If the annotations within a chosen database are above average, then an above average amount of proteins will be identified during database matching to mass spectra. However, if the database is very well annotated such as that of model organisms e.g. Arabidopsis (The Arabidopsis Genome Initiative, 2000) and rice (International Rice Genome Sequencing Project, 2005), then an abundance of proteins will have both quantitative and qualitative information (Champagne and Boutry, 2013). This was prevalent in the MU datasets (Tables S3.1 and S3.2), which was searched against the all monocotyledonous plants database (*Liliopsida*) available from UniProtKB, using reviewed sequences only, where proteins identified and found to be differentially regulated, were largely annotated with more protein descriptions.

Because tef is considered to be a non-model crop species whose genome has only been recently sequenced (Cannarozzi *et al.*, 2014), the amount of annotated information therein cannot compare to that of model plant organisms. It is important to note that the tef genome, transcriptome and proteome has only been moderately-annotated and thus would consequently lead to not all tef proteins being identified during database searching (shown in Tables 3.1 to 3.4). Nevertheless, a significant amount of proteins within the TE and TEU datasets do contain protein annotations and therefore can be used to make protein inferences through bioinformatics analyses (further discussed in Chapter 4) while those unidentified proteins may lead to discovery of some unique new targets within the tef genome.

3.3.4 OrthoMCL database search and Venn diagram generation

Due to the sheer volume of data generated by splitting the results from one iTRAQ experiment into 3 different datasets, a decision of which dataset to pursue for further interpretation using bioinformatics approaches had to be made. To achieve this and compare differentially regulated proteins within each of the three datasets, a Venn diagram of orthologous protein groups was predicted using OrthoMCL (Fischer *et al.*, 2011) in Figure 3.6. The tool was used to find groups of proteins that were common between each of the differentially regulated datasets. Because each dataset contained proteins identified and statistically analysed from the same iTRAQ experiment but matched to two different databases (Tef Extended and *Liliopsida*), the protein identifiers (protein IDs or accessions) were different and OrthoMCL was used to group them together based on protein sequence.

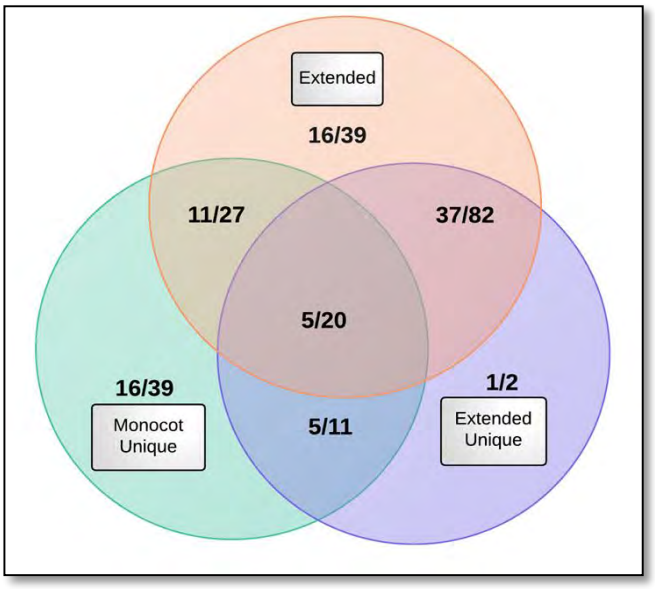

Fig. 3.6 Venn diagram (asymmetrical) of proteins and group distributions, of the TE (Extended), TEU (Extended unique) and MU (Monocot Unique) differentially regulated datasets classified using OrthoMCL tools. The first number within circles indicates the number of OrthoMCL designated protein groups achieved through BLASTP searches and second indicates the number of proteins that are present in all of those groups.

As seen in Figure 3.6, the three datasets have differing amounts of overlap. The TE and MU differentially regulated datasets each have 16 orthologous groups and 39 proteins unique to those datasets while only 1 orthologous group and 2 proteins are solely present in the TEU differentially regulated dataset. The largest amount of overlap is seen between the TE and TEU differentially regulated datasets (37 orthologous groups and 87 proteins). Because the TE and TEU differentially regulated datasets are essentially the same, the only difference being peptides uniquely classified to proteins in one dataset (TEU) and peptides classified as both unique and non-unique to proteins in the other dataset (TE), the large overlap is expected. Furthermore, 5 orthologous groups and 11 proteins are shared between the TEU and MU differentially regulated datasets and 11 orthologous groups and 27 proteins are shared between the MU and TE differentially regulated datasets. In addition a total of 5 orthologous groups and 20 proteins were commonly shared between all three differentially regulated datasets (Fig. 3.6), establishing the usefulness of searching more than one database.

Because a large overlap is observed between the TE and TEU differentially regulated datasets, it can be established that the TE differentially regulated dataset is well represented, containing proteins identified through the use of both unique and non-unique peptide mass spectra scans that meet FDR thresholds. In addition, the TE dataset is a more comprehensive differentially regulated dataset (foreground), containing 211 proteins with a number of spliced variants, significantly changing in response to dehydration stress while the TEU dataset contains approximately half the amount (111 proteins, with no occurrences of spliced variants) shown to be statistically significant. Because iTRAQ experiments on the whole do not usually produce large amounts of peptide reads per protein (Karp *et al.*, 2010), the use and manipulation

of only uniquely scanned peptides for protein identification has been shown to drastically limit the volume of confident proteins identified in the study (Gupta and Pevzner, 2009; Cappadona *et al.*, 2012). This is especially prevalent by the marginal difference observed in the amount of proteins identified between the TE and TEU differentially regulated datasets, 211 and 111 proteins, respectively.

One could then argue that perhaps it is better to use a 'cross-species identification' approach for non-model plant systems, where a generic (non-specific plant species) but well-annotated database is used for protein identification (Carpentier *et al.*, 2008a; 2008b; Romero-Rodriguez *et al.*, 2014), as in the case with the MU dataset (Tables S3.1 and S3.2). In this dataset, 174 proteins were found to be differentially regulated and were generated using only uniquely scanned peptides during database searching and furthermore contained more proteins with usable descriptions and annotations for bioinformatics inference. Although this approach is widely-used for non-model plant systems (Carpentier *et al.*, 2008a; 2008b) such as tef and many others (Hajheidari *et al.*, 2005; Carpentier *et al.*, 2007), using the same approach is not ideal as the number and confidence of identified proteins is reduced (Romero-Rodriguez *et al.*, 2014). This was shown by the amount of proteins identified by use of the MU database (4328 proteins in total of which 174 were differentially regulated) and the use of the TE database (5727 proteins in total of which 211 were differentially regulated). The difference in the total amount of proteins detected can be explained by the fact that either some species-specific proteins will not be present during cross-species identification or those homologous proteins that are present will show small evolutionary differences in their sequences (Romero-Rodriguez *et al.*, 2014). Thus, the use of a very specific but moderately-annotated database (the TE database), would detect more proteins present, highlight more proteo-bioinformatics changes that are unique to the organism under study, and also improve annotation and curation within the existing tef database.

3.4 Brief Conclusion

The iTRAQ mass spectrometry technique coupled to peptide OFFGEL fractionation and appropriate database searching with the Tef Extended and *Liliopsida* databases was used to generate three database matched protein datasets. These datasets, TE, TEU and MU each contained a large amount of database matched proteins when using the software tool PEAKS Studio 6.0. A total of 5727 proteins were identified for the TE dataset, 2656 proteins identified with the TEU dataset and 4328 proteins identified with the MU dataset. Furthermore, data refinement tools and statistical analysis through the use of the R-package, protViz, allowed differential protein expression, whereby 211 proteins for the TE dataset, 111 proteins for the TEU dataset and 174 proteins for the MU dataset were found to be differentially regulated in response to dehydration stress. A reciprocal BLAST search through the use of OrthoMCL with all three datasets was able to display common proteins and protein groups as well as show the overlap between the three differentially regulated datasets.

From these results, it was established that the TE differentially regulated dataset is well-represented with usable protein descriptions and annotations. Furthermore, the amount of proteins shown to be differentially regulated (211 in total, encompassing a fair amount of unique peptides and spliced variants) can be used for bioinformatics analyses (discussed in Chapter 4) and to make valid inferences to tef proteomic dehydration stress response (discussed in Chapter 4). In addition, some of the theoretical protein identities found through iTRAQ analysis (Table 3.1, Table 3.3 and Table S3.1) will be validated, to observe if a biological response is indeed present in tef with imposed dehydration stress (discussed in Chapter 5).

3.5 Supplementary Material

Table S3.1 MU high-abundance proteins in response to dehydration stress. Protein ID: protein identifier; protein description: identified protein; hydrated: averaged quantitative expression values belonging to hydrated labels 115-117 (a); dehydrated: averaged quantitative expression values belonging to dehydrated labels 118-121 (b); fold change: change in quantitative expression between hydrated and dehydrated values (b/a), where values > 1 display an increase in fold change; p-value: associated p-value for statistical significance (p-value ≤ 0.05).

Protein ID	Protein description	Hydrated-a	Dehydrated-b	Fold change	p-value
A2Y6W7	30S ribosomal protein S1	6.33	7.3	1.15	0.001
H6WCP2	glutamine synthetase cytosolic isozyme 1-2	5.67	6.92	1.22	0.002
I1R1T6	---NA---	5.93	7.07	1.19	0.002
M0SLX3	probable nadh dehydrogenase fqr1-like 2	1	5.81	5.81	0.002
M0TFQ1	phosphoglucose isomerase	7.89	8.44	1.07	0.003
J3L4S9	---NA---	8.3	8.84	1.07	0.004
J3NA21	probable lrr receptor-like serine threonine-protein kinase	4.94	6.35	1.29	0.004
K3Z4Y8	serine threonine-protein kinase	6.11	6.94	1.14	0.004
F2EE28	chaperonin cpn60	1	5.27	5.27	0.006
B8YAB7	---NA---	1	5.47	5.47	0.007
K7UL12	ribulose bisphosphate carboxylase	6.59	7.66	1.16	0.007
M0S163	spermidine synthase 1 short	8.17	8.67	1.06	0.007
I1PWQ2	bisphosphoglycerate-independent phosphoglycerate mutase	6.7	7.52	1.12	0.008
Q09EN6	atp synthase subunit mitochondrial	8.23	8.69	1.06	0.008
C5WRI0	phospholipase a1	1	5	5.00	0.009
C5XWE5	glycerophosphodiester phosphodiesterase gdpdl3	9.76	9.95	1.02	0.009
M8AZH0	enolase 1 a	10.8	10.95	1.01	0.009
K3XM31	60S ribosomal protein l13-2	8.96	9.35	1.04	0.01
K3YEI2	v-type proton atpase 16 kda proteolipid subunit	7.19	7.66	1.07	0.01
M7ZWS2	red chlorophyll catabolite reductase	6.92	7.54	1.09	0.01
I1P1A2	udp-glucuronate:xylan alpha-glucuronosyltransferase 1	7.34	7.84	1.07	0.011
K4APG4	abc transporter i family member	1	5.25	5.25	0.011
K7V3Y4	bisphosphoglycerate-independent phosphoglycerate mutase	7.96	8.17	1.03	0.011
M8B2N0	peroxidase 3 -rare cold-inducible protein	8.33	8.53	1.02	0.011
C5XWJ8	elongation factor ef-tu	1	5.44	5.44	0.012
M8BP49	atp synthase f1 sector subunit alpha	1	5.38	5.38	0.012
B8BH08	---NA---	6.21	7.17	1.15	0.013
J3LQZ7	---NA---	6.83	7.58	1.11	0.013
B5AMJ8	alpha-glucan h isozyme	7.82	8.51	1.09	0.015
Q2EZ09	probable lrr receptor-like serine threonine-protein kinase	5.56	6.74	1.21	0.015
I1HPZ8	cytosolic isocitrate dehydrogenase	5.03	6.77	1.35	0.016
F2D6T1	4-diphosphocytidyl-2-c-methyl-d-erythritol	8.28	8.62	1.04	0.017
I1PYX0	---NA---	8.99	9.27	1.03	0.017
J3MU66	heat shock protein 81-1	8.52	9.12	1.07	0.017

Q688K0	guanosine nucleotide diphosphate dissociation inhibitor 1	6.51	7.07	1.09	0.017
Q6TUC6	glycine hydroxymethyltransferase	6.11	6.54	1.07	0.017
J3LZ24	glyceraldehyde-3-phosphate dehydrogenase	7.57	7.85	1.04	0.018
H6T2S0	dna-directed rna polymerase subunit gamma	7.66	8.19	1.07	0.019
Q8M914	atp synthase subunit chloroplastic a	6.41	7.58	1.18	0.019
B4FQ90	vacuolar-processing enzyme	7.08	7.65	1.08	0.02
J3LEH7	glyceraldehyde-3-phosphate dehydrogenase	6.6	7.02	1.06	0.02
K4A8N1	serine methylase flags	7.65	7.98	1.04	0.021
C5YBL4	heat shock protein 82	7.67	8.35	1.09	0.022
I1Q436	homeobox protein knotted-1-like 11	9.04	9.32	1.03	0.022
J3KZC5	s-adenosylmethionine synthase 2	7.83	8.32	1.06	0.022
K3Y5D4	elongation factor g	8.37	8.68	1.04	0.022
K3XG60	protein iq-domain 14	7.92	8.43	1.06	0.023
F2E4B4	elongation factor 2 -ef-2	5.68	6.8	1.20	0.025
Q84TA3	leukotriene a-4 hydrolase homolog	6.44	7.13	1.11	0.025
A3BY14	mitochondrial outer membrane protein porin 1	7.69	8.31	1.08	0.027
M0SYI2	f-box protein skip23	8.37	8.86	1.06	0.028
C5WYH9	disease resistance rpp13-like protein 1	7.22	7.65	1.06	0.029
F2CTM7	proline-trna ligase	8.1	8.37	1.03	0.029
K3YQN8	luminal-binding protein 3	10.32	10.5	1.02	0.031
K3Z4G5	heat shock cognate 70 kda protein 2	10.76	10.9	1.01	0.032
Q5SBT2	atp synthase subunit	6.37	7.19	1.13	0.032
K3XUT7	arf guanine-nucleotide exchange factor gnom	4.86	6.29	1.29	0.033
B8A2B4	guanine nucleotide-binding protein subunit beta-like protein a	5.08	7.12	1.40	0.034
K7TM76	serine threonine-protein kinase	1	5.13	5.13	0.034
M0TC33	cytochrome p450 714b2	5.79	7.17	1.24	0.037
M7Z031	splicing factor u2af large subunit a	6.13	6.98	1.14	0.037
M7ZWJ0	mixed-linked glucan synthase 2 a	7.68	8.04	1.05	0.037
F1DI22	phosphoenolpyruvate carboxylase	9.17	9.76	1.06	0.038
J3NDJ0	40S ribosomal protein S9-2	9.81	9.85	1.00	0.038
K3ZIK0	ribulose bisphosphate carboxylase oxygenase	7.43	7.81	1.05	0.038
K3ZRY2	sodium proton antiporter	5.86	6.64	1.13	0.038
M0U0C4	serine protease	7.23	7.63	1.06	0.038
N1QQ42	probable rna helicase sde3	5.07	6.37	1.26	0.04
B6SYK7	---NA---	7.05	7.69	1.09	0.041
K3YJA1	14-3-3-like protein gf14-c	9.5	9.75	1.03	0.041
K7UU03	probable histone h2a variant 3	6.74	7.32	1.09	0.041
Q6Q297	probable 4-coumarate-ligase 3	6.33	7.08	1.12	0.041
J3NCI1	tropinone reductase-like 2	6.38	7.09	1.11	0.042
M7ZF75	alpha-l-arabinofuranosidase 1	6.83	7.86	1.15	0.042
C0PP08	---NA---	5.98	7.03	1.18	0.043
J3L4R5	---NA---	8.42	8.93	1.06	0.043
Q5JK51	---NA---	7.27	7.89	1.09	0.043

Q0D574	probable indole-3-acetic acid-amido synthetase	6.82	7.63	1.12	0.045
B6TV07	glutamate decarboxylase	6.76	7.45	1.10	0.046
E9KJF9	atp synthase subunit	6.78	7.83	1.15	0.046
K4AD01	cholesterol dehydrogenase	7.66	8.21	1.07	0.046
M8C7G4	cell division cycle protein 48 homolog	9.99	10.25	1.03	0.047
K7AF06	14-3-3-like protein gf14-d	5.34	6.84	1.28	0.048
J3MGQ3	atp-dependent zinc metalloprotease ftsh	6.98	7.86	1.13	0.049
Q94HT9	transposon tf2-12 polyprotein	1	5.07	5.07	0.05

Table S3.2 MU low-abundance proteins in response to dehydration stress. Protein ID: protein identifier; protein description: identified protein; hydrated: averaged quantitative expression values belonging to hydrated treatment labels 115-117 (a); dehydrated: averaged quantitative expression values belonging to dehydrated treatment labels 118-121 (b); fold change: change in quantitative expression between hydrated and dehydrated values (b/a), where values < 1 display a decrease in fold change; p-value: associated p-value for statistical significance (p-value ≤ 0.05).

Protein ID	**Protein description**	**Hydrated-a**	**Dehydrated-b**	**Fold change**	**p-value**
M0RYP4	plasma membrane atpase	8.97	7.98	0.89	0.002
A2YVN3	probable lrr receptor-like serine threonine-protein kinase	7.16	6.49	0.91	0.003
I1GKL4	chromosome region maintenance 1 protein homolog	5.48	1	0.18	0.004
J3LTP8	60S ribosomal protein Sl4-1	4.66	1	0.21	0.004
J7F224	50S ribosomal protein	6.12	1	0.16	0.004
K3ZXX1	40S ribosomal protein S12	5.7	1	0.18	0.004
K7V3R8	ribonucleoprotein complex subunit 1-like protein 1	7.4	6.14	0.83	0.004
Q40693	heat shock cognate 70 kda protein 2	5.57	1	0.18	0.004
A7DX42	phosphoenolpyruvate carboxylase 1	8.1	7.29	0.90	0.005
I1HY72	---NA---	6.14	1	0.16	0.005
K3YHA5	glutamate decarboxylase	4.78	1	0.21	0.005
Q2QM69	2-methyl-6-phytyl-hydroquinone methyltransferase	9.13	8.84	0.97	0.006
H6UDT9	light-harvesting complex i 11 kda protein	5.3	1	0.19	0.008
K3XID0	---NA---	4.91	1	0.20	0.01
K3Z686	protease do-like	9.35	9.1	0.97	0.01
Q2R2U7	transposon tf2-12 polyprotein	5.49	1	0.18	0.01
C0P5W4	electron transfer flavoprotein subunit	7.44	6.19	0.83	0.011
D6C788	phosphoenolpyruvate	5.2	1	0.19	0.011
J3KWV6	nadh-cytochrome b5 reductase-like protein	6.93	6.22	0.90	0.011
K7TWX2	---NA---	6.77	5.29	0.78	0.011
M0YEW4	---NA---	7.4	6.83	0.92	0.011
B9EUA8	dual specificity protein kinase pom1	6.98	5.74	0.82	0.012
B9F6D9	protein stichel-like 3	7	5.95	0.85	0.012
B9G402	heat stress transcription factor b-1	8.32	7.87	0.95	0.012
J3N642	signal recognition particle 54 kda	6.42	4.8	0.75	0.012
M0WSZ3	mannose-1-phosphate guanyltransferase alpha	7.65	6.98	0.91	0.012
Q0DF89	ferredoxin	5.49	1	0.18	0.012
I1HQV5	---NA---	4.93	1	0.20	0.013

B4FM07	2-cys peroxiredoxin	7.65	7.14	0.93	0.014
F4Y5B3	heat shock protein 90-1	7.34	5.96	0.81	0.014
A2X358	50S ribosomal protein	9.01	8.38	0.93	0.015
Q0IPN8	---NA---	5.24	1	0.19	0.015
C5YTE5	disease resistance rpp8-like protein 3	5.71	1	0.18	0.016
B4FAE5	uncharacterised protein	5.19	1	0.19	0.018
F2D6S1	mitochondrial substrate carrier family protein b	7.81	7	0.90	0.018
K4AI98	photosystem ii repair protein psb27-	8.66	8.42	0.97	0.018
M0U0B7	enolase	6.63	5.53	0.83	0.018
I1NU41	---NA---	5.66	1	0.18	0.019
M7Z1H8	auxin efflux carrier component 8	7.19	6.35	0.88	0.019
Q7XP71	gypsy retrotransposon integrase-like protein 1	5.97	1	0.17	0.019
I1I0K8	programmed cell death protein	6.82	5.75	0.84	0.02
K3XE49	protein translocase subunit	9.44	9.1	0.96	0.02
F2EBB5	Histone	8.22	7.72	0.94	0.021
B8AKV0	haloacid dehalogenase-like hydrolase domain-containing protein	7.35	6.98	0.95	0.022
M8CMG6	germin-like protein 1	5.07	1	0.20	0.022
C5XKF4	---NA---	6.74	5.99	0.89	0.023
I1IP77	glucose-1-phosphate adenylyltransferase	6.64	5.63	0.85	0.023
A7DX76	phosphoenolpyruvate carboxylase 3	6.47	4.96	0.77	0.024
E0WCS9	phosphoenolpyruvate carboxylase 1	7.96	7.02	0.88	0.024
K4AMW7	disease resistance rpp13-like protein 4	7.24	6.7	0.93	0.026
M0YAF5	---NA---	7.78	6.8	0.87	0.026
K3XDR1	ddt domain-containing protein ddb	7.77	7.2	0.93	0.028
B8Y2W7	30S ribosomal protein	8.99	8.52	0.95	0.029
K6Z6J1	ascorbate peroxidase cytosolic	7.41	6.7	0.90	0.029
M0SUI3	e3 ubiquitin-protein ligase	7.9	7.44	0.94	0.03
C5Z2G8	peroxidase 5	6.08	4.65	0.76	0.031
K7TZY7	chlorophyll a-b binding protein	10.37	10.22	0.99	0.032
C5YNB9	nad h azoreductase	8.99	8.71	0.97	0.036
C7J7J1	beta-galactosidase 14	9.35	8.99	0.96	0.036
I1HEZ1	probable sarcosine oxidase	6.62	5.64	0.85	0.036
C5YHN2	glycine-trna ligase	6.9	4.9	0.71	0.037
J3NAA0	disease resistance protein rpm1	8.01	7.51	0.94	0.037
K3YTT8	nadp-dependent d-sorbitol-6-phosphate dehydrogenase	6.58	5.53	0.84	0.037
B9F8C3	fumarate hydratase	9.42	5.19	0.55	0.038
K3YWM5	60S ribosomal protein l14-1	9.77	9.64	0.99	0.038
M0TZF6	elongation factor 1-gamma 2	8.29	7.9	0.95	0.038
B3SHC6	atp synthase subunit	8.04	7.53	0.94	0.039
C5WWE2	nad-dependent malic enzyme 59 kda	6.39	4.96	0.78	0.04
M8BG98	trna (cytosine-c)-methyltransferase	6.66	6.08	0.91	0.04
F8U875	glucose phosphomutase 1	7.37	6.41	0.87	0.041
I1I2I7	photosynthetic nadh subunit of sub complex b	9.16	8.88	0.97	0.041

K3YL45	translocon at the inner envelope membrane of chloroplasts 62	10.01	9.91	0.99	0.041
M0SDW7	fructose-bisphosphate cytoplasmic isozyme 1	8.99	8.37	0.93	0.042
Q8RYY4	---NA---	6.57	5.52	0.84	0.043
I1IF76	atp synthase delta	9.89	9.71	0.98	0.044
Q9SDJ2	magnesium-protoporphyrin ix monomethyl ester	9.18	9.03	0.98	0.044
F2CXC6	---NA---	8.01	7.09	0.89	0.045
M0VAF3	patellin-3	7.26	6.69	0.92	0.045
Q5MD10	glutamine synthetase root isozyme 5	10.01	9.69	0.97	0.045
Q6Z1V6	nadh-quinone oxidoreductase subunit	8.61	7.98	0.93	0.045
C5X3Z0	gamma carbonic anhydrase	7.67	7.06	0.92	0.046
M0U3I3	cysteine synthase	7.06	6.17	0.87	0.046
M0WBU8	squamosa promoter-binding-like protein 15	8.07	7.67	0.95	0.047
J3KZJ0	---NA---	8.6	7.86	0.91	0.048
J3M442	importin subunit alpha-1b	8.46	7.82	0.92	0.048
M0RV12	ribulose bisphosphate carboxylase oxygenase activase	8.95	8.55	0.96	0.048
M0S1D8	uncharacterised protein	7.69	6.72	0.87	0.049
M7ZWK6	transaminase a	8.62	8.19	0.95	0.049
K3Z133	homoserine kinase	9.83	9.63	0.98	0.05

The following supplementary files can be found in the folder "Supplementary data files-Chapter 3 iTRAQ analysis" on the online cloud storage application, Dropbox via the following link: https://www.dropbox.com/sh/u9s5mjt4xjsjup3/AAAxUCWJe0HnYkfy0sONjJZAa?dl=0

File S3.1 TE dataset (all 5727 proteins with expression values, fold changes and p-values).

File S3.2 TE - two group analysis (box-plots for all 5727 proteins).

File S3.3 TE - Sanity Check (quality check 1)

File S3.4 TE - Heat map (quality check 2)

File S4.1 TEU dataset (all 2656 proteins with expression values, fold changes and p-values).

File S4.2 TEU - two group analysis (box-plots for all 2656 proteins)

File S4.3 TEU- Sanity Check (quality check 1)

FileS4.4 TEU - Heat map (quality check 2)

File S5.1 MU dataset (all 4328 proteins with expression values, fold changes and p-values).

File S5.2 MU – two group analysis (box plots for all 4328 proteins).

File S5.3 MU - Sanity check (quality check 1)

FileS5.4 MU - Heat map (quality check 2)

Chapter 4: Bioinformatics analyses of tef proteins

4.1 Introduction

Comparative proteomics is a commonly used tool to understand plant responses to abiotic stresses (Ghosh and Xu, 2014). Such analyses allow detection of changes in protein abundance under the imposed stress conditions and correlation of those with plant phenotypic responses (Abreu *et al.*, 2013; Ghosh and Xu, 2014). By making inferences to the plant proteomic profile in response to biological stimuli, the measurement of protein abundance changes allow key proteins and biological processes to be highlighted for further investigation (Baginsky, 2009; Vanderschuren *et al.*, 2013).

For the successful identification of proteins involved in biological activities within the cell, the database used for data mining is of key importance and ultimately determines the accuracy in functional interpretation of results (Cañas *et al.*, 2006; Balbuena *et al.*, 2011). A number of open source programs are available for functional annotation of protein datasets and subsequent query for enriched Gene Ontology (GO) processes (Ashburner *et al.*, 2000). However, many of them are tailored to well-known plant organisms (Carpentier *et al.*, 2008b), with the most common being Arabidopsis (*Arabidopsis thaliana*), rice (*Oryza sativa*) and maize (*Zea mays*), due to their well-annotated, high-quality genomes that undergo constant improvement (Carpentier *et al.*, 2008b). For non-model plant species such as tef, functional annotation and GO-term analysis is more challenging and requires the use of more "data-suitable" annotation tools.

There are currently plant proteome and genome databases available online that can be used for data processing of non-model plant organisms (Jorrin-Novo *et al.*, 2009; Balbuena *et al.*, 2011). However, when searching for protein sequences from a non-model plant against databases generated from closely-related plant species based on orthologous similarity (Grossmann *et al.*, 2007; Balbuena *et al.*, 2011), proteins involved in processes that are unique to the non-model plant are often not characterised (Carpentier *et al.*, 2008a). Therefore, having a database generated from the sequenced non-model plant genome of the species under study provides more data and aids subsequent interpretation.

The recent sequencing of the tef genome (Cannarozzi *et al.*, 2014), has provided a platform for analysis of proteomic data and identification of important biological processes in tef that may be involved in response to dehydration stress. This tef genomics resource has allowed the use of an array of open source bioinformatics tools that are most suited to a recently sequenced genome to deduce protein annotation, functional classification, GO-term evaluation and enrichment analysis, as well as to observe biological pathways of interest in response to dehydration stress. In the previous chapter, a total of 211 proteins belonging to the TE dataset were found to be differentially regulated in response to dehydration stress (discussed in Chapter 3). Thus, in order to maximize the assignment of biological meaning to significant

protein identities and to determine the functional pathways in which they might be operating, the TE database was used for further functional classification and annotation of regulated proteins. The programs listed below have been used to analyse tef proteins identified using the TE database. Analysis included both foreground (differentially regulated and statistically significant 211 proteins identified in the TE dataset) and background proteins (all 5727 proteins identified in the TE dataset):

1.) MapMan Analysis, using the MapMen suite of tools consisting of Mercator (Lohse *et al.*, 2014) and MapMan (Thimm *et al.*, 2004; Usadel *et al.*, 2005), for mapping and profiling the differentially regulated tef proteins onto biological pathways or processes.
2.) Blast2GO (Conesa *et al.*, 2005; Gotz *et al.*, 2008), for protein classification, annotation and retrieval of Gene Ontology (GO) terms (Ashburner *et al.*, 2000) used in functional enrichment analysis in response to dehydration stress.
3.) KEGG: Kyoto Encyclopaedia of Genes and Genomes (Kanehisa and Goto, 2000)(Kanehisa and Goto, 2000), retrieved through Blast2GO, used for the investigation of interesting, stress responsive biological pathways.

4.2 Methods to bioinformatics evaluation

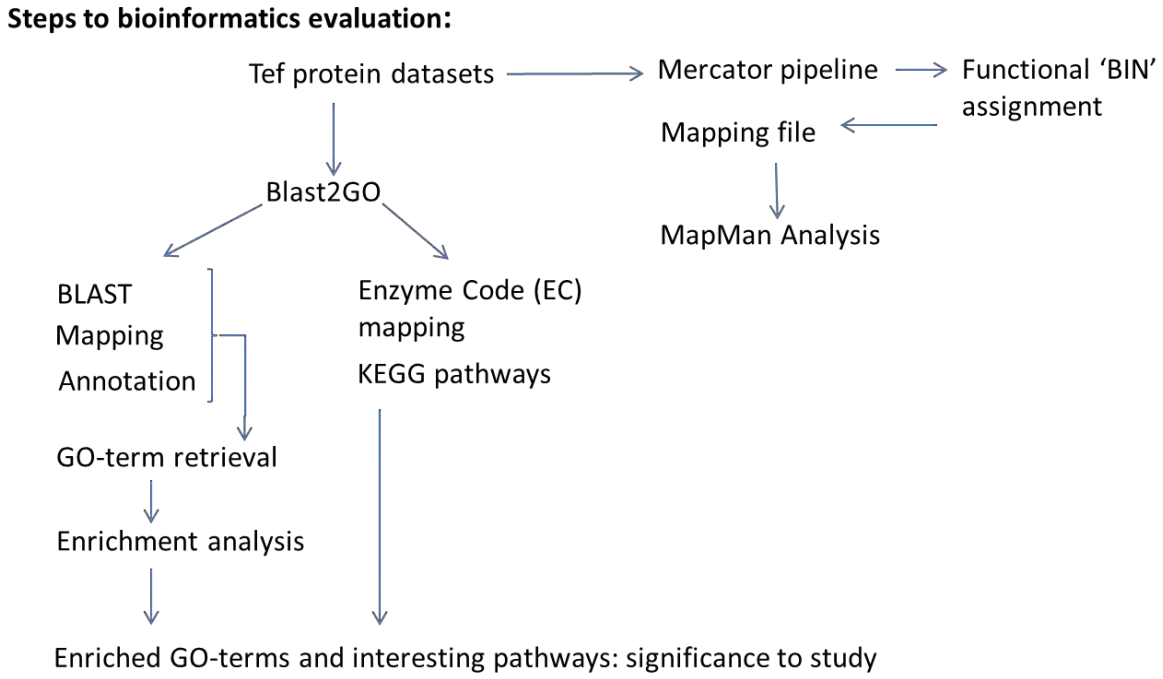

Fig. 4.1 Outline of the steps used to bioinformatically evaluate tef proteins.

4.2.1 MapMan analysis using MapMen tools: Mercator and MapMan

4.2.1.1 Mercator pipeline

Mercator (http://mapman.gabipd.org/web/guest/app/Mercator), is a web-based annotation tool used to assign functional terms to protein or nucleotide sequences by using the MapMan 'BIN' ontology and has been specifically designed for functional annotation of plant 'omics' data (Lohse *et al.*, 2014). The tef foreground was uploaded to the web-application Mercator and a mapping file consisting of functional predictions for protein sequences was generated through searching six available databases (3 BLAST-based, 2 reverse position-specific BLAST based and InterProScan, namely Arabidopsis TAIR proteins release 10, SwissProt/Uniprot plant proteins, TIGR5 rice proteins, Clusters of orthologous eukaryotic genes database-KOG and Conserved Domains Database-CDD). The Mercator output consisting of a mapping file containing the search results of each protein query computed into the most likely functional BIN was then ready for input and 'mapping' using the MapMan visualisation tool.

4.2.1.2 MapMan analysis

MapMan (Thimm *et al.*, 2004; Usadel *et al.*, 2005) software (version 3.6.0) was used for functional classification and visualisation of differentially expressed tef proteins through their 'Scavenger' and 'ImageAnnotator' modules. The Scavenger module classifies genes or proteins into a redundancy-reduced ontology in the form of BINs or sub-BINs and generates a 'mapping file', which is essentially

the output from Mercator, while the ImageAnnotator module visualises the expression data onto schematic diagrams 'maps' of biological processes or pathways based on the generated mapping file (Usadel *et al.*, 2009). A text file (.txt) of the tef foreground proteins, containing unique protein identifiers and changes in expression between hydrated (control) and dehydrated (experimental) values in the form of Log_2 fold changes and associated p-values, was imported into the experimental data files in MapMan. The experimental data file was configured to display significantly high or low-abundance proteins based on Log_2 scale with a p-value ≤ 0.05. Statistical significance according to the MapMan BIN code system, was then calculated using a Wilcoxon rank sum test which tests whether the average response of a particular BIN differs from the response of all the other remaining BINs (Usadel *et al.*, 2005). This experimental data file together with the mapping file generated by Mercator was then used to map proteins onto known biological pathways or processes in MapMan.

4.2.2 Blast2GO tools for protein annotation and functional enrichment analysis

4.2.2.1 Protein annotation and GO-term retrieval

Tef protein annotation, classification and retrieval of GO-terms was performed using the BLAST, mapping and annotation set of tools in Blast2GO version 2.8 (Conesa *et al.*, 2005; Gotz *et al.*, 2008). Both datasets (in FASTA format) consisting of tef foreground and background proteins were matched against the UNIPROTKB/SwissProt database using the BLASTP algorithm with the following parameters: report a maximum of twenty blast hits, with a blast expect value of $1e^{-3}$ and minimum high scoring segment pairs (HSPs) length equal to 33. Subsequent to initiating BLAST steps, the steps to mapping and annotation were then employed for GO-term retrieval using the Blast2GO default parameters (*E*-value filter if $1e^{-6}$, an hsp-hit coverage cut-off of 0, annotation cut-off of 55, and GO weight of 5) with the September 2014 database. To gather as much information from the tef protein sequences, InterProScan 5.0 (Quevillon *et al.*, 2005) and GO-enzyme code mapping steps were also performed using the Blast2GO default settings.

4.2.2.2 Functional enrichment analysis

Subsequent to tef protein annotation and classification, functional enrichment of GO-terms was initiated using the Fisher's exact test for statistical significance (Bluthgen *et al.*, 2005) in Blast2GO. For input, both tef foreground and background annotation files were merged as one file (.annot), which was then used as a reference set. A list of protein identifiers containing individually named contigs from the foreground was used as a test set. For enrichment of high and low-abundance proteins, the test set (foreground) was separated into two lists, one containing high-abundant protein identifiers and the other containing low-abundant protein identifiers. Fisher's exact test was employed for both high and low-abundance proteins to show both over and under-represented GO-terms. A two-sided Fisher's exact test, using a term filter of 0.05 and term filter mode set as false discovery rate (FDR) with the removal duplicate IDs, was utilized. The graph generated to better display functional enrichment of GO-terms for

tef foreground-high-abundance proteins and tef foreground-low-abundance proteins vs. tef background proteins was performed using GraphPad Prism 6.0 software by utilising the exported chart data for GO-term enrichment in Blast2GO.

Blast2GO tools were used for the MU dataset for all differentially regulated proteins (as described in Chapter 3, section 3.3.3). Bioinformatics analysis of the MU dataset was initiated to observe if having a generic well-annotated database, comprised of evolutionary closely-related plant species would provide a larger range of significantly enriched GO-terms for stress response inference. The same BLAST, mapping and annotation parameters (as described above) for both foreground (identified and differentially regulated 174 proteins) and background (all 4328 proteins identified using the *Liliopsida* database in UNIPROTKB/SwissProt) were used. Once GO-terms were retrieved, functional enrichment analysis was employed in the same manner as before. However, no significant GO-term results were achieved with the term filter set as FDR and cut-off of less than or equal to 0.05.

4.2.2.3 KEGG pathways

To observe biological pathways of interest, pathway maps from KEGG (Kanehisa and Goto, 2000) showing highlighted enzyme codes (ECs) for proteins within the TE dataset were retrieved through Blast2GO for further interpretation. The graph generated to better display the top 21 KEGG pathways was performed using GraphPad Prism 6.0 software by utilising the exported chart data for KEGG pathways in Blast2GO.

4.3 Results and Discussion

4.3.1 Mercator functional 'BIN' assignment

The program Mercator (Lohse *et al.*, 2014), was used as an annotation tool for the classification of tef protein sequences into essentially non-redundant functional BINs and sub-BINs and for the generation of a 'mapping' file to be used in MapMan (Thimm *et al.*, 2004; Usadel *et al.*, 2005). This allowed for the identification of biological processes which respond to dehydration stress. Figure 4.2 shows a pie chart depicting the proportion of tef foreground proteins allocated to each Mercator BIN. The Mercator ontology has a total of 35 BINs. A table of BIN code definitions is provided in supplementary information (Table S4.1). As seen from the pie chart (Fig. 4.2), 110 (52.34%) out of the 211 proteins identified in the tef foreground list were allocated into 34 functional BINs, while a large percentage of sequences (47.66%) were left un-assigned (BIN 35) according to the Mercator BIN code system.

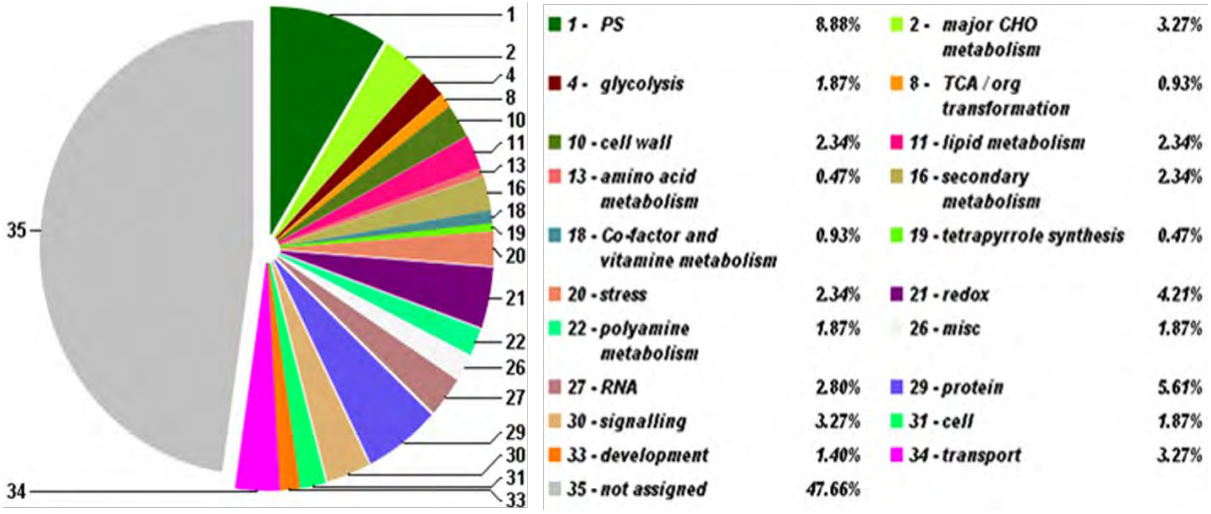

Fig. 4.2 Mercator functional 'BIN' allocations of tef foreground in response to dehydration stress. The pie-chart displayed was generated by the Mercator annotation tool to display functional BIN assignment.

4.3.2 MapMan analysis

To further visualise the distribution of differentially regulated tef foreground proteins according to Mercator mapping file, the MapMan analysis tool (Thimm *et al.*, 2004; Usadel *et al.*, 2005), was used. This analysis allowed profiling of Mercator allocated BINs containing protein sequences onto pre-set biological pathways with quantitative expression values on a Log_2 scale. An overview of the proteins allocated to each MapMan BIN is shown in Figure 4.3, with a corresponding list being presented in Table 4.1.

A total of 149 proteins, were mapped using the MapMan visualisation tool (Table 4.1) with the largest proportion of proteins (45 proteins, 30%) being unclassified (BIN 35, Fig. 4.3). BIN numbers 1, 2, 4, 8, 10, 11, 13 and 19 representing photosynthesis (19 proteins, 12%), carbohydrate metabolism (7 proteins, 4.7%), glycolysis (2 proteins, 1.3%), TCA organic acid transformation (2 proteins, 1.3%), cell wall (4

proteins, 2.6%), lipid metabolism (5 proteins, 3.6%), amino acids (1 protein, 0.7 %) and tetrapyrrole synthesis (1 protein, 0.7 %), respectively, were among the protein classes in which expression changed the least (indicated by the lightly coloured blocks in Fig. 4.3). A significant decrease in protein abundance occurred in BINs numbered 18, 22 and 34. These represented co-factor and vitamin metabolism (2 proteins, 1.3%), polyamine metabolism (4 proteins, 2.6%) and transport (7 proteins, 4.7%), respectively (red blocks, Fig. 4.3). In contrast, the BIN numbers 16, 20, 21, 26, 27 29, 30, 31 and 33, coloured by blue blocks and corresponding to secondary metabolism (3 proteins, 2%), stress response (5 proteins, 3.3%), redox (7 proteins, 4.7%), miscellaneous proteins (4 proteins, 2.6%), RNA processing (6 proteins, 4%), protein metabolism (12 proteins, 8%), signalling (7 proteins, 4.7%), cell (3 proteins, 2.5%) and development (3 proteins, 2.5%), respectively, were significantly increased in protein abundance in response to dehydration stress.

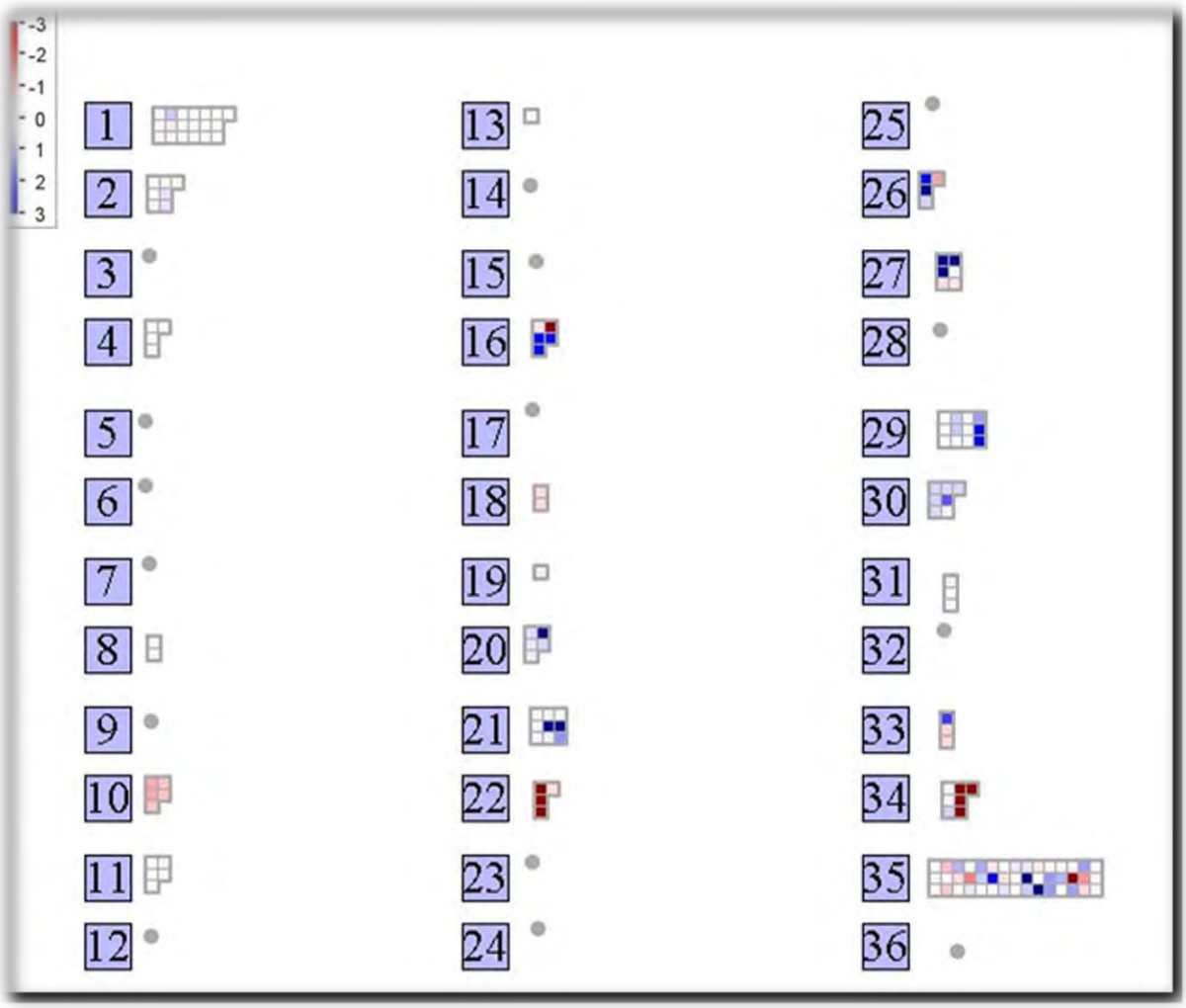

Fig. 4.3 Overview of MapMan BIN allocations for tef foreground proteins involved in different metabolic processes. A total of 149 proteins out of 211 were mapped to 21 functional BINS. Blue blocks display high-abundance proteins while red blocks display low-abundance proteins as a consequence of dehydration stress on a scale of -3 to 3. Grey circles represent BINs to which no proteins were allocated. The names of the respective BINS can be found in Table 4.1 and in the associated the text.

Table 4.1 Overview of MapMan assigned BINs for tef foreground proteins. The table displays BIN numbers, BIN names, number of elements (proteins) allocated to each BIN and associated probabilities (p-value ≤ 0.05), based on Wilcoxon Rank Sum test, where significance is given in ranks by testing if the changes within a BIN are more extreme than the changes in all the remaining proteins.

BIN	Name	Elements (no. of proteins allocated)	p-value
22	polyamine metabolism	4	0.002
10	cell wall	4	0.005
21	Redox	7	0.021
34	Transport	7	0.021
29	Protein	12	0.024
20	Stress	5	0.052
30	Signalling	7	0.078
18	Co-factor and vitamin metabolism	2	0.106
27	RNA	6	0.206
26	Misc	4	0.251
2	major CHO metabolism	7	0.454
33	Development	3	0.464
16	secondary metabolism	3	0.472
1	PS	19	0.491
19	tetrapyrrole synthesis	1	0.575
4	Glycolysis	2	0.707
8	TCA / org transformation	2	0.718
13	amino acid metabolism	1	0.740
11	lipid metabolism	5	0.749
31	Cell	3	0.946
35	not assigned	45	0.981
Total amount of proteins mapped in tef foreground		**149**	

A large number of differentially regulated and statistically significant proteins were placed in the 'not assigned' BIN (BIN 35), resulting in considerable underrepresentation within pathway 'maps'. The MapMan visualisation tool ultimately relies on the 'mapping' file (BIN allocation) processed in Mercator which from our result (Fig. 4.2), is poorly annotated and does not truly reflect all the proteins present in the tef foreground dataset. Although the displayed MapMan generated pathway maps do not fully represent all the proteins present in the tef foreground list, those that have been successfully mapped to BIN classes can still be used to substantiate bioinformatics findings obtained using alternative processing tools.

More biological information on pathways could have been obtained if a pre-set mapping file (present in MapMan) of a closely related plant, such as *Sorghum bicolor* (Cannarozzi *et al.*, 2014), was used. However, this would require searching tef sequences for orthologous proteins that correspond to the chosen plant species and subsequent conversion of protein identifiers for accurate mapping and visualisation. While this approach has been used by many researchers as a valid method of bioinformatics inference (Rotter *et al.*, 2009; Hiremath *et al.*, 2011; Nestler *et al.*, 2011; Abdalla and Rafudeen, 2012; Wang *et al.*, 2013), we used the available tef genome and transcriptome (Cannarozzi *et al.*, 2014) in our bioinformatics approach to minimise false identification and provide further information to the tef

proteome. This approach has been successfully used by Balbuena *et al.* (2012) for large scale characterisation of the rhizome proteome from *Equisetum hyemale* (horsetail), where the sequenced transcriptome was translated to a concatenated protein search database and used in further bioinformatics analyses.

4.3.3 Blast2GO

4.3.3.1 Enrichment Analysis (Fisher's exact test)

In order to observe functional GO-term enrichment of differentially regulated tef proteins in response to dehydration stress (Chapter 3, section 3.3.3), a Fisher's exact test for statistical significance in Blast2GO (Bluthgen *et al.*, 2005; Conesa *et al.*, 2005) was used. Table 4.2 gives enriched GO-terms for both high and low-abundance proteins. A total of 50 GO-term processes were functionally enriched, of which 22 GO-terms were found for high-abundance proteins (Table 4.2, A) and 28 GO-terms were found for low-abundance proteins (Table 4.2, B). All of these belonged to the classification categories (ontologies) of cellular component (CC), molecular function (MF) and biological process (BP).

To summarise the findings and have a more visual representation of functionally enriched GO-terms, terms were filtered and reduced to the most specific annotations (most specific GO, using FDR as a term filter and term filter value of less than 0.05) and represented as histograms for both high and low-abundance proteins (see Figure 4.4 A, B, respectively). The enriched GO-terms shown were reduced to 29 most specific terms in total of which 11 GO-terms were found for high-abundance proteins (Fig. 4.4A) and 18 GO-terms for low-abundance proteins (Fig. 4.4B) in the classification categories CC, MF and BP. The reduced and most-specific GO-terms are shown in more detail in Table S4.2, provided in the supplementary figures.

In terms of high-abundance proteins, all enriched GO-term processes were shown to be over-represented in response to dehydration stress (Table 4.2, A; Fig. 4.4A), including a minority of GO-terms that were not necessarily linked to plant systems. These would include: negative regulation of neurogenesis (GO:0050768) and negative regulation of axonogenesis (GO:0050771) in the category BP. These biological processes are linked to the generation of new neurons and developing axons from stem cells after foetal and post-natal development has been completed (Kornblum, 2007; Zhu *et al.*, 2010). In the category CC, the only GO-term M band (GO:0031430), which relates to the location of specific proteins that link thick filaments in a sarcomere of muscle tissue (Smith, 2000), was represented. This particular GO-term has been linked to six glycolytic enzymes that catalyse successive reactions along the glycolytic pathway in *Drosophila melanogaster* (fruit fly) flight muscle (Sullivan *et al.*, 2003). These include fructose-1, 6-bisphophate aldolase, glycerol-3-phosphate dehydrogenase, glyceraldehyde-3-phosphate dehydrogenase, triose phosphate isomerase, phosphoglycerate kinase and phosphoglycerol mutase. While there is an obvious evolutionary gap between the fruit fly and tef, their metabolic processes with regards

to glycolysis are similar, implying that the analysis of GO-terms is often a method of showing common gene ontological processes, independent of the species under study (Primmer *et al.*, 2013).

Table 4.2 Functional enrichment analysis of GO-terms allocated to proteins differentially expressed by either displaying high-abundance (A) or low-abundance (B) in response to dehydration stress. Significant GO-terms retrieved through Fisher's exact test in Blast2GO. Enrichment analysis of GO-terms have been selected on the basis of False discovery rate (FDR < 0.05) and p-value (p-value < 0.01). GO-ID: the ID number of the GO-term. Term: description of the GO-term. Ontology: GO-terms categorization in cellular component (CC), molecular function (MF) biological processes (BP) groups. FDR: the proportion of false positives that was determined by calculating the false discovery rate corresponding to each enrichment score. P-value: p-value indicating the statistical significant difference between the fraction of proteins assigned to GO-term retrieved and the fraction of all proteins within the background set assigned to the same GO-term. Over/under: GO terms under or over-represented in the test set in response to dehydration stress.

GO-ID	Term	Ontology	FDR	P-value	Over/Under
A: High-abundance					
GO:0031430	M band	CC	4.06E-2	8.76E-5	Over
GO:0016656	monodehydroascorbate reductase (NADH) activity	MF	1.40E-3	1.37E-7	Over
GO:0005093	Rab GDP-dissociation inhibitor activity	MF	1.55E-3	3.59E-7	Over
GO:0005097	Rab GTPase activator activity	MF	1.55E-3	1.06E-6	Over
GO:0005092	GDP-dissociation inhibitor activity	MF	1.55E-3	1.06E-6	Over
GO:0090315	negative regulation of protein targeting to membrane	BP	1.55E-3	1.06E-6	Over
GO:0090313	regulation of protein targeting to membrane	BP	1.55E-3	1.06E-6	Over
GO:0050771	negative regulation of axonogenesis	BP	1.55E-3	1.06E-6	Over
GO:0032851	positive regulation of Rab GTPase activity	BP	2.08E-3	2.45E-6	Over
GO:0032483	regulation of Rab protein signal transduction	BP	2.08E-3	2.45E-6	Over
GO:0032313	regulation of Rab GTPase activity	BP	2.08E-3	2.45E-6	Over
GO:0050768	negative regulation of neurogenesis	BP	2.08E-3	2.45E-6	Over
GO:0010721	negative regulation of cell development	BP	2.08E-3	2.45E-6	Over
GO:0043903	regulation of symbiosis, encompassing mutualism through parasitism	BP	4.26E-3	5.42E-6	Over
GO:0031345	negative regulation of cell projection organization	BP	6.27E-3	8.60E-6	Over
GO:0019321	pentose metabolic process	BP	2.55E-2	3.76E-5	Over
GO:0048227	plasma membrane to endosome transport	BP	2.96E-2	4.64E-5	Over
GO:0009610	response to symbiotic fungus	BP	3.11E-2	5.74E-5	Over
GO:0009608	response to symbiont	BP	3.11E-2	5.74E-5	Over
GO:0090317	negative regulation of intracellular protein transport	BP	3.11E-2	6.42E-5	Over
GO:0032387	negative regulation of intracellular transport	BP	3.11E-2	6.42E-5	Over
GO:0051224	negative regulation of protein transport	BP	3.11E-2	6.42E-5	Over

B: Low-abundance					
GO:0010598	NAD(P)H dehydrogenase complex (plastoquinone)	CC	5.06E-4	1.48E-7	Over
GO:0030076	light-harvesting complex	CC	1.16E-2	2.16E-5	Over
GO:0005576	extracellular region	CC	1.21E-2	2.47E-5	under
GO:0051139	metal ion: proton antiporter activity	MF	1.93E-3	1.13E-6	Over
GO:0010486	manganese: proton antiporter activity	MF	1.93E-3	1.13E-6	Over
GO:0005384	manganese ion transmembrane transporter activity	MF	2.67E-3	2.61E-6	Over
GO:0008810	cellulase activity	MF	3.47E-3	4.76E-6	Over
GO:0046524	sucrose-phosphate synthase activity	MF	6.24E-3	9.18E-6	Over
GO:0048038	quinone binding	MF	9.07E-3	1.51E-5	Over
GO:0015491	cation: cation antiporter activity	MF	9.07E-3	1.51E-5	Over
GO:0015299	solute: proton antiporter activity	MF	1.61E-2	3.47E-5	Over
GO:0015298	solute: cation antiporter activity	MF	1.61E-2	3.47E-5	Over
GO:0016168	chlorophyll binding	MF	3.11E-2	7.94E-5	Over
GO:0010258	NADH dehydrogenase complex (plastoquinone) assembly	BP	7.53E-5	7.37E-9	Over
GO:0010257	NADH dehydrogenase complex assembly	BP	1.30E-4	2.54E-8	Over
GO:0044248	cellular catabolic process	BP	1.93E-3	1.13E-6	under
GO:0009768	photosynthesis, light harvesting in photosystem I	BP	2.67E-3	2.07E-6	Over
GO:0055071	manganese ion homeostasis	BP	2.67E-3	2.61E-6	Over
GO:0030026	cellular manganese ion homeostasis	BP	2.67E-3	2.61E-6	Over
GO:0051275	beta-glucan catabolic process	BP	3.47E-3	4.76E-6	Over
GO:0042547	cell wall modification involved in multidimensional cell growth	BP	3.47E-3	4.76E-6	Over
GO:0030245	cellulose catabolic process	BP	3.47E-3	4.76E-6	Over
GO:0071421	manganese ion transmembrane transport	BP	1.06E-2	1.88E-5	Over
GO:0006828	manganese ion transport	BP	2.06E-2	4.65E-5	Over
GO:0044712	single-organism catabolic process	BP	2.54E-2	5.97E-5	under
GO:0010042	response to manganese ion	BP	2.79E-2	6.84E-5	Over
GO:0009117	nucleotide metabolic process	BP	4.39E-2	1.20E-4	under
GO:0006753	nucleoside phosphate metabolic process	BP	4.39E-2	1.20E-4	under

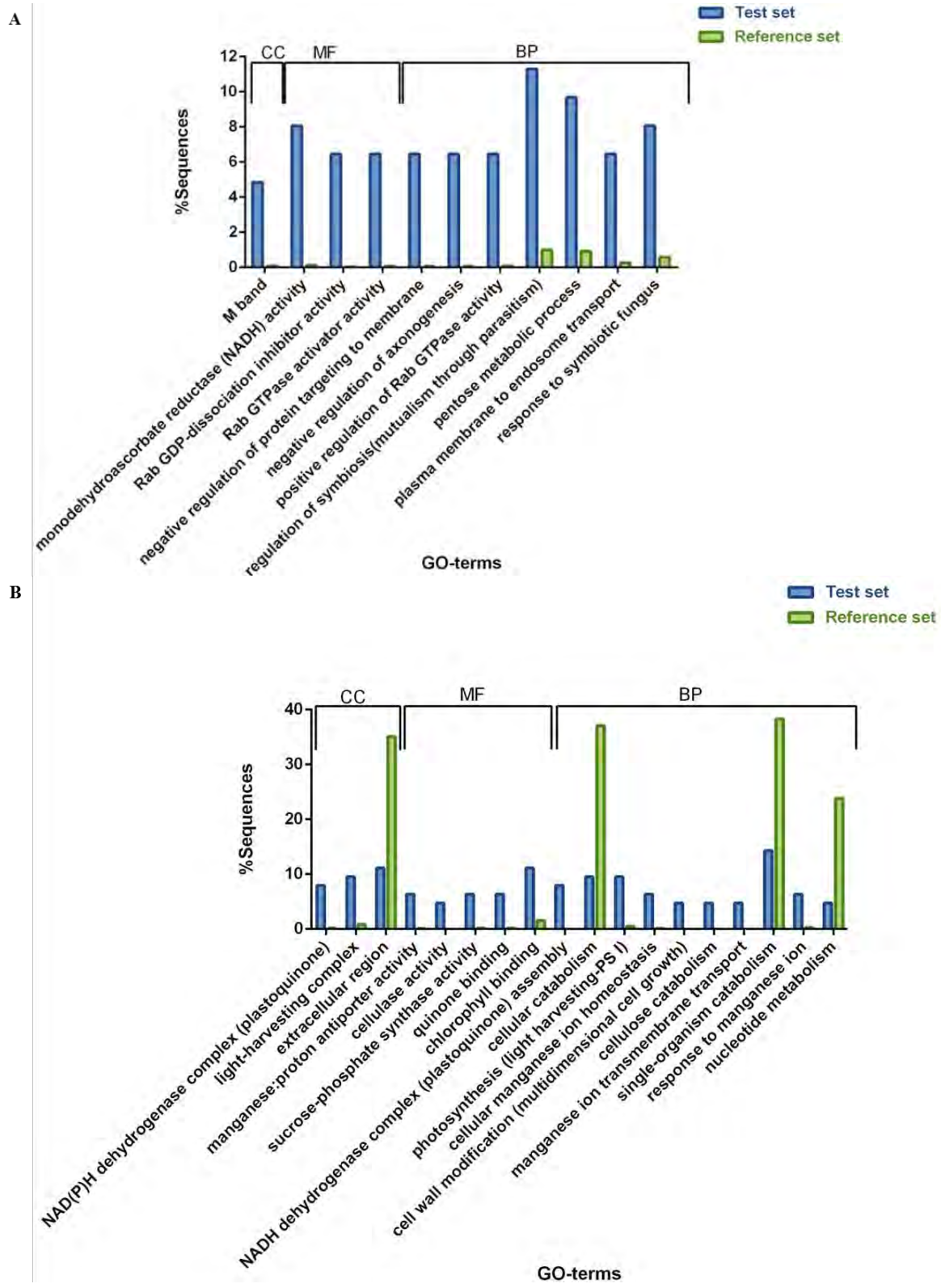

Fig. 4.4 GO-term classifications of tef differentially regulated proteins in response to dehydration stress. Histograms show tef high-abundance proteins (test set) vs. tef background proteins (reference set) (A) and tef low-abundance proteins (test set) vs. tef-background proteins (reference set) (B). GO-term results are summarised into three main categories: cellular component (CC), molecular function (MF) and biological process (BP). The y-axis designates the percentage of protein sequences in each GO-term classification while the x-axis displays the GO-term classifications. GO-terms have been reduced to the most specific terms using a term filter mode of FDR and term filter value of less than 0.05. Results are shown as coloured bars (blue: test set and green: reference set).

In the category MF, monodehydroascorbate reductase (NADH) activity (GO:0016656), was depicted as the most significant GO-term (Table 4.2, A; FDR = 1.40E-3, p-value = 1.37E-7) in response to dehydration stress. This is further highlighted when represented in the histogram showing the most specific GO-terms (Fig. 4.4, A; 8.1% protein sequences). Monodehydroascorbate reductase (MDHAR, EC 1.6.5.4), is one of the key enzymes involved in ascorbate reduction (Morell *et al.*, 1997) and functions in reducing the oxidised form of ascorbate (monodehydroascorbate) before being returned to the ascorbate pool (Morell *et al.*, 1997; Asada, 2006; Huang *et al.*, 2013; Shin *et al.*, 2013). Monodehydroascorbate has been proposed to be an indicator of oxidative stress within plant tissues, playing an important role in cellular response against accumulating ROS due to increasing stress conditions (Gill and Tuteja, 2010; Huang *et al.*, 2013; Shin *et al.*, 2013).

The Rab family of cellular processes active in the regulation of vesicular membrane traffic (Zarsky *et al.*, 1997) were equally over-represented in response to dehydration stress, under the category MF (Table 4.2, A; Fig. 4.4A). These include: GTPase activator activity (GO:0005097), Rab GDP-dissociation and GDP-dissociation inhibitor activity (GO:0005093, GO:0005092). Other GO-terms involved in membrane trafficking, usually in response to some type of intracellular signalling (Cheung and De Vries, 2008) such as regulation and negative regulation of protein targeting to membrane and intracellular protein transport (GO:0090313, GO:0090315, GO:0090317), were equally over-represented, in response to dehydration stress (Table 4.2, A; Fig. 4.4A), under the category BP. The flow of membrane constituents between endomembrane structures and the plasmalemma is critical for the maintenance of cellular homeostasis in response to signal transduction (Chrispeels *et al.*, 1999; Cheung and De Vries, 2008). Similarly, transport processes such as plasma membrane to endosome transport (GO:0048227) and other inhibitory transport processes such as the negative regulation of intracellular transport and negative regulation of protein transport (GO:0032387, GO:0051224), were over-represented in response to dehydration stress (Table 4.2, A; Fig. 4.4A).

Co-incidentally the GO-terms allocated to biological processes responsible for regulating membrane trafficking and the flow of proteins and other macromolecules to numerous endpoints inside and outside the cell through a signalling cascade (Vernoud *et al.*, 2003; Cheung and De Vries, 2008), were over-represented in response to dehydration stress. These include, the regulation of Rab GTPase activity (GO:0032313), regulation of Rab protein signal transduction (GO:0032483) and positive regulation of Rab GTPase activity (GO:0032851) (Table 4.2, A; Fig. 4.4A). The Rab family of small GTP-binding proteins function as molecular alterations that cycle between ‘active’ and ‘inactive’ states within the cell through the binding and hydrolysis of GTP (Vernoud *et al.*, 2003), thereby controlling the endocytic network in plants (Agarwal *et al.*, 2008). Interestingly, the stress-inducible small GTP-binding protein Rab7 gene (*PgRab7*) isolated from *Pennisetum glaucum*, a relatively drought-stress tolerant food grain crop grown in India, has been reported to increase tolerance to abiotic stresses such as drought and

increased salinity in transgenic tobacco (Agarwal *et al.*, 2008). Similarly, the Rab7 gene (*TaRab7*) isolated from wheat leaves infected with the wheat stripe rust pathogen (*Puccinia striiformis* f. sp. *tritici*), was proposed to play an important role in early stages of wheat-stripe rust fungus interaction and stress tolerance (Liu *et al.*, 2012).

During dehydration stress, tef responses to biotic challenges such as fungal or bacterial infections are also important, as the GO-terms, response to symbiont and symbiotic fungus (GO:0009608, GO:0009610) and regulation of symbiosis, encompassing mutualism through parasitism (GO:0043903), were highly over-represented in the BP category (Fig. 4.4A; 8.1 and 11.3% protein sequences, respectively). Although tef has been proposed to be relatively free from damage by insects or competition from weeds (Stallkneecht *et al.*, 1993), at least 22 species of fungi and 3 pathogenic nematodes have been previously associated with tef (Bekele, 1985; Stallkneecht *et al.*, 1993), substantiating our GO-term findings.

The GO-term pentose metabolic process (GO:0019321), in the category BP, was significantly over-represented in response to dehydration stress (Fig. 4.4A; 9.7% protein sequences). The pentose phosphate pathway has been reported to having a dual role in oxidative stress response in plants (Couee *et al.*, 2006). Firstly, by providing an available source of soluble-sugars that can either be involved in ROS-producing metabolic pathways (Couee *et al.*, 2006; Jain, 2013) or alternatively, by being involved in the active production of NADPH, a major co-factor required in the antioxidant ascorbate-glutahione cycle (Couee *et al.*, 2006; Gill and Tuteja, 2010). In addition, these soluble sugars have been proposed to act as nutrient and metabolite signalling molecules that activate specific signalling pathways leading to imperative gene modification and proteomic changes in response to a number of stresses (Couee *et al.*, 2006).

A substantial amount of GO-terms were enriched in low-abundance proteins (Table 4.2, B), where a number of processes in response to dehydration stress were both over and under-represented in the ontology categories CC, MF and BP (Fig. 4.4B). Amongst the under-represented GO-terms, where reference set GO-terms were significantly higher than test set GO-terms, most were involved in cellular catabolic processes (Table 4.2, B; Fig. 4.4B). These would include: extracellular region (GO:0005576), cellular catabolic process (GO:0044248), single-organism catabolic process (GO:0044712) and nucleotide metabolism (GO:0009117, GO:0006753), all of which were represented by their large percentages of reference set GO-terms (Fig. 4.4B; 35%, 37.1%, 38.3% and 23.8% protein sequences, respectively).

The functional enrichment of GO-terms found to be over-represented in low-abundance proteins (Table 4.2, B), were commonly linked to quinone cycling in the plastoquinone pool during oxidative phosphorylation, namely quinone binding (GO:0048038), NADH dehydrogenase complex

(plastoquinone) assembly (GO:0010258, GO:0010257) and NAD(P)H dehydrogenase complex (plastoquinone) (GO:0010598), in the categories MF, CC and BP, respectively (Table 4.2, B; Fig. 4.4B). The complexes NADH dehydrogenase and NAD(P)H dehydrogenase, both function in reducing plastoquinones during the flow of electrons when ATP is generated (Keunen *et al.*, 2011; Jacoby *et al.*, 2012). While NADH dehydrogenase functions in cellular respiration in the mitochondria (Keunen *et al.*, 2011), NAD(P)H dehydrogenase is localised in the thylakoid membranes of chloroplasts, participating in cyclic electron transport reactions around photosystem I and chlororespiration (interactions linking respiratory electron transport chain and photosynthetic electron transport chain in thylakoid membranes of chloroplasts) (Peltier and Cournac, 2002; Peng *et al.*, 2011; Eugeni Piller *et al.*, 2011). NAD(P)H, in particular, has been proposed to lessen oxidative stress in plants (Peng *et al.*, 2011).

Increased supplying of ATP for photosynthesis has been reported during environmental stress conditions, particularly during drought stress (Rumeau *et al.*, 2005). However, since photosynthetic metabolism under water-deficit stress is reported to be responsible for the production of large amounts of free radicals (Gill and Tuteja, 2010), these processes in effect, are decreased in tef in an attempt perhaps to minimise ROS production. In further support that reduced ROS production is important in the tef dehydration stress response, GO-terms involved in photosynthetic processes, such as photosynthesis, light harvesting complex (GO:0030076), light harvesting in photosystem I (GO:0009768) and chlorophyll binding (GO:0016168), were over-represented in low-abundance proteins as well, in the categories CC, BP and MF, respectively (Table 4.2, B; Fig. 4.4B). In addition, the GO-terms linked to ROS-producing processes through the generation of additional ATP, such as the transfusion of solutes in the form of cations and protons across membranes (GO:0015491, GO:0015299, GO:0015298), were over-represented in the category MF, in response to dehydration stress (Table 4.2, B; Fig. 4.4B).

The categories, transport and response of metal ions, were well over-represented in low-abundance proteins, by being commonly placed in the ontology categories MF and BP, in response to dehydration stress (Table 4.2, B; Fig. 4.4B). These would include the GO-terms metal ion: proton antiporter activity (GO:0051139), manganese: proton antiporter activity (GO:0010486), manganese ion transmembrane transporter activity (GO:0005384), manganese ion homeostasis (GO:0055071), cellular manganese ion homeostasis (GO:0030026), response to manganese ion (GO:0010042) and manganese ion transport, transmembrane transport (GO:0006828, GO:0071421). The positively charged micronutrient, manganese, is required during the splitting of water in photosystem II, when photosynthesis occurs (Peiter *et al.*, 2007; Huda *et al.*, 2013) and has been reported to play important roles as a co-factor and activator of enzymes in various sub-cellular compartments (Chrispeels *et al.*, 1999; Huda *et al.*, 2013), particularly in manganese superoxide dismutase (Mn-SOD) antioxidant activity in plant mitochondria (Keunen *et al.*, 2011).

To avoid toxicity within cell tissues, cytosolic manganese concentrations need to be kept low (Huda *et al.*, 2013) and are usually transported out of the cytosol by metal transporters (Peiter *et al.*, 2007), where they are either localised to the plant cell membrane or to the vacuolar membrane where metals are sequestered into large moderately inert compartments (Peiter *et al.*, 2007). If manganese concentrations are not carefully monitored in plant cells, toxicity is usually indicated by chlorosis, brown specks, necrosis and crinkled leaves (Peiter *et al.*, 2007), which arise due to the inhibition of chlorophyll synthesis (Peiter *et al.*, 2007). The disruption of manganese ion transport and homeostasis and consequent decreased protein abundance in tef, comes as no surprise in response to dehydration stress as photosynthetic potential has been shown to decrease at water contents below 55% RWC (Chapter 2, section 2.3.3). The decrease in photosynthesis and inhibition of chlorophyll synthesis, would ultimately lead to increased manganese concentrations and toxicity within tef plant cells due to metal transport and cellular manganese homeostasis disruption.

Potential modification of the cell wall, particularly in the form of the terms cellulase activity (GO:0008810), cellulose catabolism (GO:0030245), beta-glucan catabolism (GO:0051275), as well as cell wall modification involved in multidimensional cell growth (GO:0042547), was also observed as over-represented in low-abundance proteins in the categories MF and BP in response to dehydration stress (Table 4.2, B; Fig. 4.4B). The effect of cell wall re-structuring and modification during stress conditions is a common phenomenon in plant cells as a consequence of turgor loss during dehydration stress (Marshall and Dumbroff, 1999; Moore *et al.*, 2006; 2008). Many plants curtail the growth of their stems and leaves when subjected to low water potential (Wu and Cosgrove, 2000) and continue to elongate the root tissues for deeper soil penetration and water mining as a result of adapting to drought conditions (Wu and Cosgrove, 2000; Moore *et al.*, 2008). In addition to seeking out extra water sources by root tissues, the cell wall is either tightened or loosened in certain tissues in response to drought conditions, resulting in the loosening of growth 'ready' tissues (increased cell wall extensibility), such as in the apices of root tissues and tightening (made inextensible) in tissues not functioning in water-uptake such as in stems and leaves (Wu and Cosgrove, 2000; Moore *et al.*, 2008).

Previous observations in tef with regards to increased primary root lengths and decreased shoot growth in response to drought conditions have been reported (Degu and Fujimura, 2010) and have been proposed to be an adaptive morphological response of tef in water-limiting environments (Degu and Fujimura, 2010). Further ultra-structural observations in tef during different stages of dehydration have shown cell wall folding and eventually breakage occurring during the latter stages of dehydration at approximately 20% RWC (Chapter 2, Fig. 2.6C, D). This could be due in part to the cell wall made inextensible or tightened as a consequence of turgor loss and the decrease in abundance of GO-term processes involved in cell wall breakdown, cellulose catabolism and cell wall modification pertaining to multidimensional growth (Fig. 4.4B).

Lastly, the GO-term sucrose-phosphate synthase activity (GO:0046524), in the category MF, was over-represented in low-abundance proteins in response to dehydration stress (Table 4.2, B; Fig. 4.4B). Sucrose phosphate synthase (EC 2.4.1.14) plays an important role in the synthesis of sucrose using substrates derived from glycolysis such as fructose-6-phosphate and UDP-glucose (Whittaker *et al.*, 2007). In correlation to being functionally enriched in low-abundance tef proteins (Fig. 4.4B), the enzyme has been previously shown to decrease in activity in the leaves of other C4 species as well, such as maize (Pelleschi *et al.*, 1997) and sugarcane (Du *et al.*, 1999), in response to dehydration stress (Whittaker *et al.*, 2007). The decline in sucrose accumulation has been proposed to be due to the decline in readily available photosynthetic triose phosphate which ultimately leads to a decline in the enzyme activity of sucrose phosphate synthase (Whittaker *et al.*, 2007).

4.3.3.2 KEGG pathways

To best describe the identification of tef proteins and enzymes involved in biological pathways active in response to dehydration stress, a histogram displaying the number of protein sequences mapped to reference canonical biological pathways from KEGG-Kyoto Encyclopaedia of Genes and Genomes (Kanehisa and Goto, 2000), using Blast2GO was generated (Fig. 4.5). A total of 3438 protein sequences were mapped to 121 KEGG pathways (Supplementary Table S4.3). The top 21 KEGG biological pathways and the number of enzymes involved in each pathway are displayed below (Fig. 4.5). Only pathways with more than 50 mapped protein sequences were analysed in depth. The largest number of protein sequences mapped, were found in the pathway "carbon fixation in photosynthetic organisms'' (174 sequences), followed by glycolysis/gluconeogenesis (129 sequences), purine metabolism and glyoxylate and dicarboxylate metabolism (128 sequences) as well as pyruvate metabolism (120 sequences). The other highly represented pathways include starch and sucrose metabolism and glutathione metabolism, each represented by 93 mapped protein sequences (Fig. 4.5).

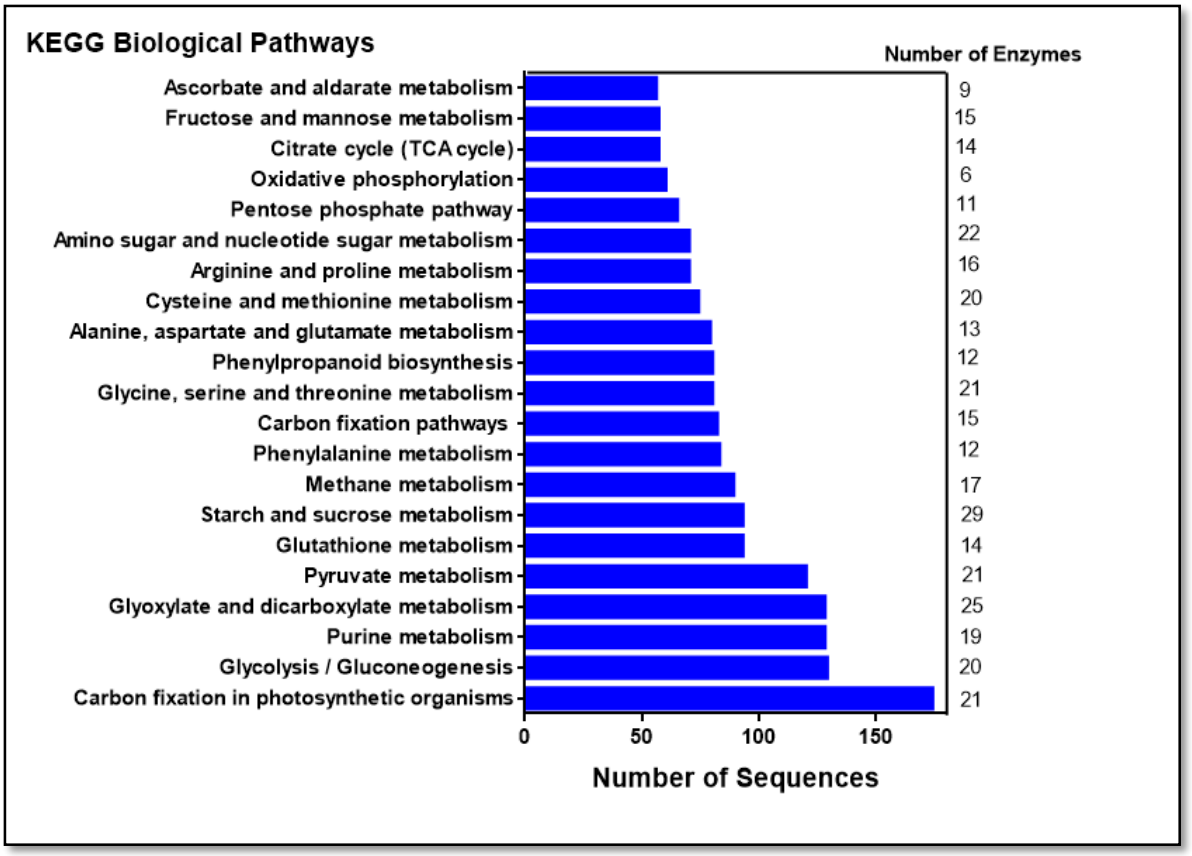

Fig. 4.5 Histogram of the top 21 biological pathways assigned to tef proteins by KEGG (retrieved through Blast2GO). The y-axis displays the name of the KEGG biological pathway and x-axis indicates the number of protein sequences allocated to each biological pathway. The number of enzymes belonging to the TE dataset and involved in each pathway have been manually added to the histogram on the right. Only pathways with more than 50 allocated protein sequences were analysed in depth.

Although a number of pathways were shown to contain proteins and enzymes belonging to the TE dataset (Fig. 4.5), only pathways known to play a pivotal role in plant stress response are discussed further. These include pathways displaying glutathione metabolism (Fig. 4.6A) as well as ascorbate and aldarate metabolism (Fig. 4.6B) each containing 14 and 9 identified enzymes, respectively (Table. 4.3). For the purpose of relevance, only the enzymes highlighted by coloured blocks and identified in the TE dataset will be described further. The full biological pathways for both glutathione and ascorbate and aldarate metabolism can be found in supplementary figures (Fig. S4.1 and Fig. S4.2, respectively).

Table 4.3 List of identified enzymes with enzyme names and codes from corresponding pathways in KEGG, glutathione metabolism (A) and ascorbate and aldarate metabolism (B).

Pathway	Enzyme	Enzyme ID
A- Glutathione metabolism	thioredoxin peroxidase	ec:1.11.1.15
	glutathione peroxidase	ec:1.11.1.12
	peroxidase	ec:1.11.1.11
	ligase	ec:6.3.2.2
	glutamyl transpeptidase	ec:2.3.2.2
	synthase	ec:2.5.1.22
	transferase	ec:2.5.1.18
	synthase	ec:2.5.1.16
	dehydrogenase (ascorbate)	ec:1.8.5.1
	reductase	ec:1.8.1.7
	peroxidase	ec:1.11.1.9
	dehydrogenase (NADP+-dependent, decarboxylating)	ec:1.1.1.44
	dehydrogenase (NADP+)	ec:1.1.1.42
	hydrolase	ec:3.4.19.13
B- Ascorbate and aldarate metabolism	peroxidase	ec:1.11.1.11
	3,5-epimerase	ec:5.1.3.18
	oxygenase	ec:1.13.99.1
	1-naphthol glucuronyltransferase	ec:2.4.1.17
	dehydrogenase (ascorbate)	ec:1.8.5.1
	reductase (NADH)	ec:1.6.5.4
	6-dehydrogenase	ec:1.1.1.22
	reductase	ec:1.1.1.19
	dehydrogenase (NAD+)	ec:1.2.1.3

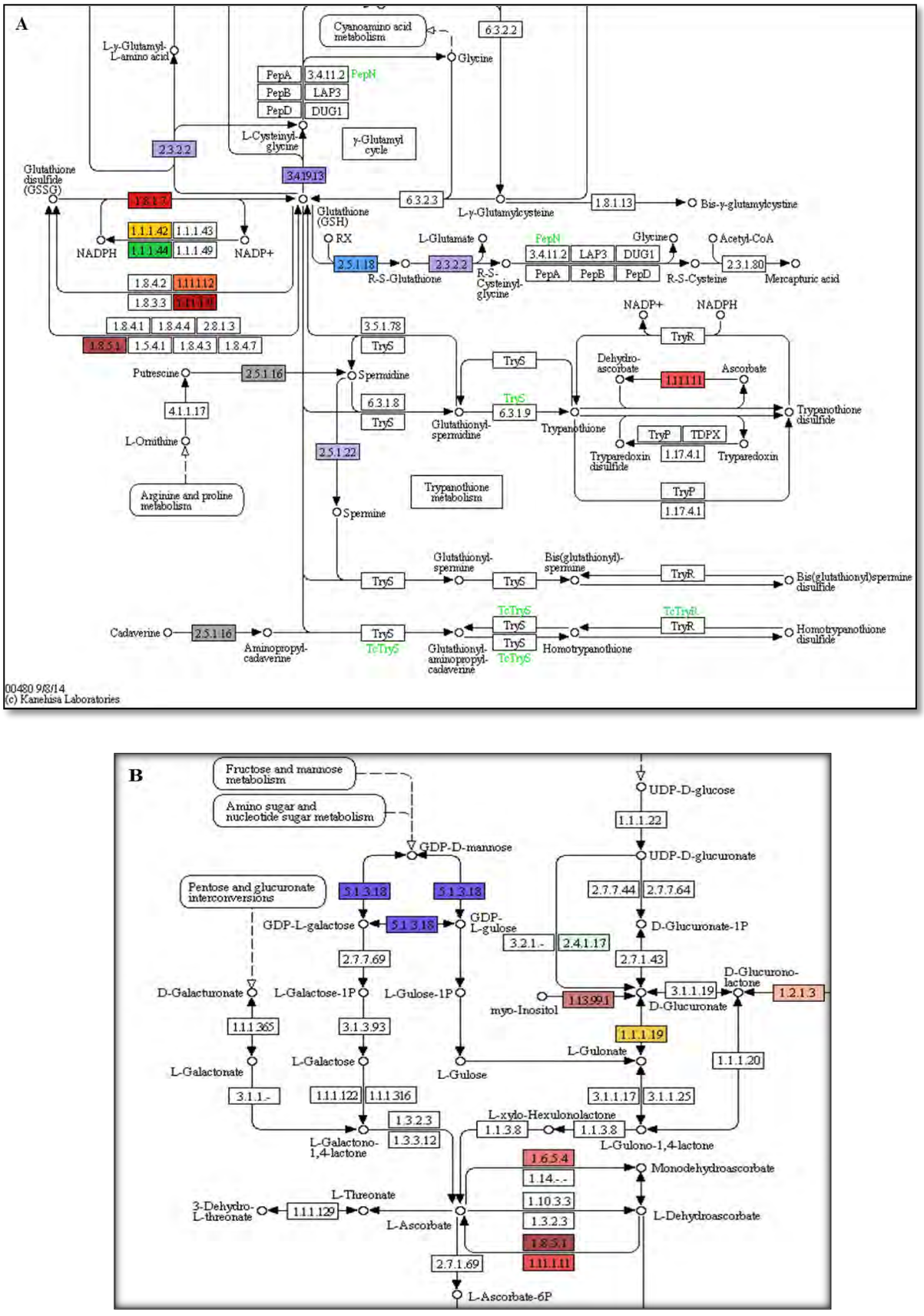

Fig. 4.6 KEGG pathways displaying glutathione metabolism (A) and ascorbate and aldarate metabolism (B). The enzymes identified from TE dataset are shown by coloured blocks, with relevant enzyme code (EC) numbers.

4.3.3.2.1 Glutathione metabolism

Glutathione is the most abundant form of organic sulphur in plants (apart from those found in amino acids) and is mostly found in its reduced form (GSH) under non-stressed conditions, with only a small proportion of it being in its oxidised state (GSSG) (Noctor and Foyer, 1998; Dixon *et al.*, 1999). A total of 14 enzymes were identified in the glutathione metabolism pathway (Table 4.3, A; Fig. 4.6, A). The reduced form of glutathione (GSH) is formed when L-glutamate and L-cysteine come together through a ligase enzyme (ec:6.3.2.2) to form ɣ-glutamylcysteine followed by the formation of GSH through the addition of glycine to the C-terminal end (Noctor *et al.*, 1998). GSH is then converted to the amino acids, L-cysteinyl-glycine and L-ɣ-glutamyl-L-amino acid by the enzymes hydrolase (ec:3.4.19.13) and glutamyl transpeptidase (ec:2.3.2.2), respectively (Table 4.3, A; Fig. 4.6, A), for use in the cyanoamino acid pathway. The enzyme, glutathione transferase (GST, ec:2.5.1.18) acts on GSH to form R-S-glutathione, which is subsequently converted to R-S-cysteinyl-glycine by the enzyme glutamyl transpeptidase (ec:2.3.2.2), releasing L-glutamate as a by-product (Table 4.3, A; Fig. 4.6, A).

GSH is oxidised to GSSG and vice-versa, through several enzymes (shown in Fig 4.6A), firstly by ascorbate dehydrogenase or dehydroascorbate reductase (ec:1.8.5.1), which is a reversible reaction, resulting in the reduced form of glutathione (GSH). Secondly GSH is converted to GSSG by glutathione peroxidase (ec:1.11.1.12) and peroxidase (ec:1.11.1.9) and thirdly GSSG is converted to GSH by the most common reaction during ascorbate-glutathione stress response, by glutathione reductase (ec:1.8.1.7), resulting in the conversion of NADPH to $NADP^+$. The co-factor $NADP^+$ can then be recycled to its reduced form NADPH, by the dehydrogenase ($NADP^+$; $NADP^+$-dependent, decarboxylating) enzymes (ec:1.1.1.42; and ec:1.1.1.4, respectively) (Table 4.3, A; Fig. 4.6, A). In addition, GSH is converted to trypanothione by a number of enzymes in the arginine and proline metabolism and trypanothione metabolism pathways, through the synthase enzymes (ec:2.5.1.22 and ec:2.5.1.16, respectively). The substrate trypanothione, is then acted upon by the enzyme ascorbate peroxidase (ec:1.11.1.11), to form the product trypanothione disulphide, resulting in the conversion of dehydroascorbate to ascorbate and efficiently linking the glutathione and ascorbate pathways (Table 4.3, A; Fig. 4.6, A).

4.3.3.2.2 Ascorbate and aldarate metabolism

A total of 9 enzymes belonging to the TE dataset were identified in the ascorbate and aldarate metabolism pathway (Table 4.3, B). Several pathways contribute to the synthesis of ascorbate (Fig. 4.6, B). The product from the pathways fructose and mannose metabolism and amino sugar and nucleotide sugar metabolism, GDP-D-mannose is acted upon by the enzyme 3, 5-epimerase (ec:5.1.3.18), in a reversible reaction to form either GDP-L-galactose and GDP-L-gulose (Fig. 4.6B; Table 4.3, B). It is the latter compound that goes on to form L-gulonate, and eventually to L-ascorbate, through a series of enzymatic reactions. Secondly, the product from amino sugar and nucleotide sugar metabolism pathway UDP-D-

glucose, is converted to UDP-D-glucuronate by the enzyme UDP-glucose 6-dehydrogenase (ec:1.1.1.22), which is then acted upon by 1-naphthol glucuronyltransferase (ec:2.4.1.17), to form D-glucuronate (Fig. 4.6B; Table 4.3, B). D-glucuronate, is additionally formed from the conversion of myo-Inositol and D-glucarate by the enzymes oxygenase (ec:1.13.99.1) and dehydrogenase (NAD^+) (ec:1.2.1.3), respectively (Fig. 4.6B; Table 4.3, B). The product, L-gulonate is then formed by the action of glucuronate reductase (ec:1.1.1.19) on D-glucuronate. The compound L-gulonate is then converted to L-ascorbate which then plays an important role in the ascorbate pool.

Ascorbate is regenerated and maintained through the ascorbate redox system, consisting of L-ascorbate, monodehydroascorbate and dehydroascorbate (Noctor and Foyer, 1998), when ascorbate is oxidised to monodehydroascorbate by ascorbate peroxidase (ec:1.11.1.11) during the decomposition of the free radical, hydrogen peroxide (Sharma and Dubey, 2005). The enzymatic reactions of monodehydroascorbate reductase (NADH) (ec:1.6.5.4), reducing monodehydroascorbate to L-ascorbate and dehydroascorbate reductase (ec:1.8.5.1), reducing non-enzymatically converted dehydroascorbate to L-ascorbate, then occurs (Fig. 4.6B; Table 4.3, B). This reaction is performed in the presence of GSH as a reducing agent which is oxidised to GSSG, linking the ascorbate and glutathione pathways in the ascorbate-glutathione cycle (Sharma and Dubey, 2005; Huang *et al.*, 2013).

4.3.3.2.3 Role of ascorbate and glutathione in plant stress tolerance

Together the ascorbate and glutathione pathways function in being powerful scavengers to the most damaging forms of ROS in response to environmental perturbations, such as the superoxide radical (O_2^-), hydroxyl free radical (OH), singlet ($_1O_2$) oxygen, hydrogen peroxide (H_2O_2) and dismutates H_2O_2 (Sharma and Dubey, 2005; Saruhan *et al.*, 2009). In addition to being antioxidants, glutathione and ascorbate have other functions in cellular metabolism. Glutathione is involved in, among others: stress signalling, in response to changes in the extracellular environment (Dixon *et al.*, 1999; Kranner *et al.*, 2006), as a pre-cursor for phytochelatins, by binding high concentrations of heavy metals, such as cadmium (Noctor *et al.*, 1998; Ha *et al.*, 1999), as part of a network regulating defence gene expression (Noctor *et al.*, 1998; Grene, 2002) and is also thought to be the major cellular redox buffer due to its redox-active thiol group (Noctor *et al.*, 1998; Kranner *et al.*, 2006; Kamies *et al.*, 2010).

Ascorbate has also been perceived to act as a signalling molecule in plant stress response and furthermore, acts as an indicator of oxidative stress levels (Zhang *et al.*, 2008; Huang *et al.*, 2013). In addition, ascorbate has been linked to photosynthetic light harvesting (Noctor and Foyer, 1998), as a substrate for cell wall peroxidases and has been proposed to assist in cell wall lignification (Saruhan *et al.*, 2009). More importantly, increased ascorbate levels have been associated with abiotic stress response, particularly with regards to drought in crops such as rice (Sharma and Dubey, 2005), maize (Chugh *et al.*, 2011) and wheat (Chakraborty and Pradhan, 2012).

In further support of the active roles of the ascorbate and glutathione pathways, our previously conducted MapMan analysis of the tef foreground proteins (see section 4.3.2) also highlighted these pathways in response to dehydration stress (Fig. 4.7).

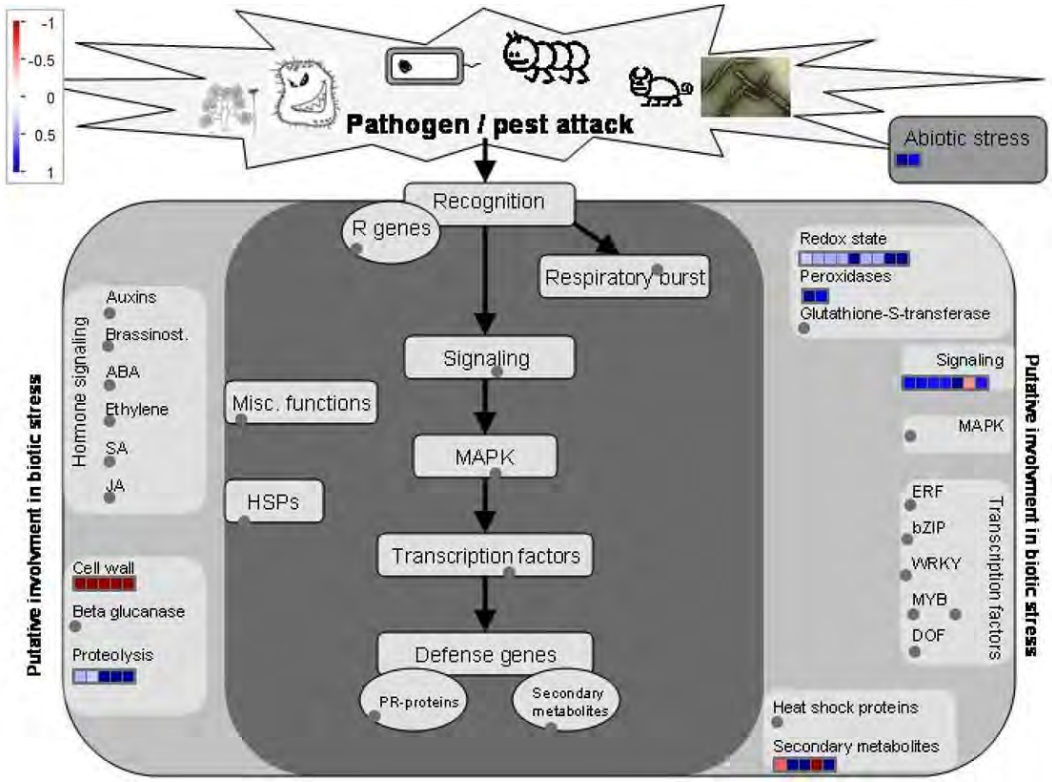

Fig. 4.7 Tef foreground proteins active in biotic and abiotic stress response. A total of 35 proteins were mapped to functional BINs related to stress response. Blue blocks display high-abundance proteins while red blocks display low-abundance proteins as a consequence of dehydration stress on a scale of -1 to 1. Grey circles represent BINs to which no proteins were allocated to.

These would include the increase in abundance of proteins involved in the pathways involved in redox state of the cell, specifically peroxidases (BIN 20 and 21) and glutathione-s-transferases (BIN 21), as well as proteins playing a pivotal role in stress response signalling (BIN 30). Additionally, proteins involved in the cell wall (BIN 10), such as cell wall synthesis and degradation were decreased in abundance in response to dehydration, supporting our previously observed over-representation of GO-terms related to cell wall catabolism and modification (GO:0008810; GO:0030245; GO:0051275; GO:0042547), in functional enrichment of low-abundance proteins (Table 4.2, B; Fig. 4.4B). The pathway of protein degradation via ubiquitination (BIN 29) was significantly increased in abundance in response to dehydration stress, while proteins related to secondary metabolites (BIN16) such as the secondary metabolism of sulphur-containing glucosinolates, isoprenoids, flavonoids and tocopherol biosynthesis were differentially expressed in response to dehydration stress (Fig. 4.7). The synthesis of the secondary metabolite tocopherol in particular, has been linked to ascorbate and glutathione utilisation in cells, whereby ascorbate acts as a secondary antioxidant, reducing the oxidised form of α-tocopherol in hydrophobic environments (Noctor and Foyer, 1998; Szarka *et al.*, 2012).

4.4 Brief conclusion

In this chapter, the TE dataset, consisting of both foreground and background proteins generated from previous iTRAQ analysis (Chapter 3), were analysed using various bioinformatics tools most suited to a non-model crop plant system. These would include use of the programs: Mercator (Lohse *et al.*, 2014), MapMan (Thimm *et al.*, 2004; Usadel *et al.*, 2005), Blast2GO (Conesa *et al.*, 2005; Gotz *et al.*, 2008) and KEGG (Kanehisa and Goto, 2000) in an attempt to retrieve as much ontological information for changing tef proteins in response to dehydration stress. The tools Mercator and MapMan used in combination were able to map and profile approximately half of the differentially regulated proteins onto known biological pathways or processes. While these programs provided some useful information for bioinformatics inference, its full potential as an annotation tool was not achieved due to many proteins being unassigned according to MapMan BIN allocations.

Blast2GO was then used for protein classification, functional annotation and retrieval of GO-terms for use in functional enrichment analysis. A total of 50 GO-terms belonging to the classification ontologies CC, MF and BP were found to be significantly over or under-represented for proteins changing in response to dehydration stress. In general, the GO-terms involved in biotic and abiotic stress response, signalling, transport, cellular homeostasis and pentose metabolic processes were enriched in high-abundance proteins. While GO-terms linked to photosynthesis and light harvesting reactions (ROS forming processes), cell wall catabolism, manganese transport and homeostasis, the synthesis of sugars and cell wall modification related to multidimensional growth were enriched in low-abundance proteins in response to dehydration stress. Lastly, KEGG was used to observe tef proteins and enzymes mapped to biological pathways, of which the stress responsive pathways, glutathione metabolism and ascorbate and aldarate metabolism were further investigated.

Based on the observed iTRAQ results (Chapter 3, section 3.3.3) and further bioinformatics interpretations gained from this chapter, a subset of proteins would need to be validated to ascertain a biological response occurring in tef with imposed dehydration stress. Because, an overall subtle shift in the total proteome is observed with dehydration stress where proteins functioning in various crucial plant maintenance processes and stress response are highlighted, the validation of proteins in the form of physiological assays and immunodetection by western blotting will be conducted on proteins functioning in maintaining tef viability and neutralising excessive ROS during dehydration stress (discussed in Chapter 5).

4.5 Supplementary Material

Table S4.1. BIN code definitions according to MapMan visualisation tool.

BIN code	Name	BIN code	Name
1	PS - photosystem	**21**	Redox
2	major CHO metabolism	**22**	polyamine metabolism
3	minor CHO metabolism	**23**	nucleotide metabolism
4	Glycolysis	**24**	biodegradation of Xenobiotcs
5	Fermentation	**25**	C1-metabolism
6	gluconeogenesis/glyoxylate cycle	**26**	miscellaneous
7	OPP	**27**	RNA
8	TCA / org transformation	**28**	DNA
9	mitochondrial electron transport/ATP synthesis	**29**	protein
10	cell wall	**30**	signalling
11	lipid metabolism	**31**	Cell
12	N-metabolism	**32**	micro RNA, natural antisense
13	amino acid metabolism	**33**	development
14	S-assimilation	**34**	Transport
15	metal handling	**35**	not assigned
16	secondary metabolism	**36**	mineral nutrition
17	hormone metabolism		
18	Co-factor and vitamin metabolism		
19	tetrapyrrole synthesis		
20	Stress		

Table S4.2. Functional enrichment analysis of GO-terms allocated to proteins differentially expressed by either displaying high-abundance (A) or low-abundance (B) in response to dehydration stress. The GO-terms displayed have been reduced to the most specific term (most specific GO) that lies above a user-defined cut-off value (threshold). Enrichment analysis of GO-terms have been selected on the basis of False discovery rate (FDR < 0.05) and p-value (p-value < 0.01). GO-ID: the ID number of the GO-term. Ontology: GO-terms categorization in cellular component (CC), molecular function (MF) biological processes (BP) groups. Term: description of the GO-term. FDR: the proportion of false positives was determined by calculating the false discovery rate corresponding to each enrichment score. P-value: p-value indicating the statistical significance of the difference between the fraction of proteins assigned to GO-term retrieved and the fraction of all proteins within the background set assigned to the same GO-term. Over/under: GO terms under or over-represented in the test set in response to dehydration stress.

GO-ID	Term	Ontology	FDR	P-value	Over/Under
A: High-abundance					
GO:0031430	M band	CC	4.06E-2	8.76E-5	over
GO:0016656	monodehydroascorbate reductase (NADH) activity	MF	1.40E-3	1.37E-7	over
GO:0005093	Rab GDP-dissociation inhibitor activity	MF	1.55E-3	3.59E-7	over
GO:0005097	Rab GTPase activator activity	MF	1.55E-3	1.06E-6	over
GO:0090315	negative regulation of protein targeting to membrane	BP	1.55E-3	1.06E-6	over
GO:0050771	negative regulation of axonogenesis	BP	1.55E-3	1.06E-6	over
GO:0032851	positive regulation of Rab GTPase activity	BP	2.08E-3	2.45E-6	over
GO:0043903	regulation of symbiosis, encompassing mutualism through parasitism	BP	4.26E-3	5.42E-6	over
GO:0019321	pentose metabolic process	BP	2.55E-2	3.76E-5	over
GO:0048227	plasma membrane to endosome transport	BP	2.96E-2	4.64E-5	over
GO:0009610	response to symbiotic fungus	BP	3.11E-2	5.74E-5	over
B: Low-abundance					
GO:0010598	NAD(P)H dehydrogenase complex (plastoquinone)	CC	5.06E-4	1.48E-7	over
GO:0030076	light-harvesting complex	CC	1.16E-2	2.16E-5	over
GO:0005576	extracellular region	CC	1.26E-2	2.47E-5	under
GO:0010486	manganese: proton antiporter activity	MF	1.93E-3	1.13E-6	over
GO:0008810	cellulase activity	MF	3.47E-3	4.76E-6	over
GO:0046524	sucrose-phosphate synthase activity	MF	6.24E-3	9.18E-6	over
GO:0048038	quinone binding	MF	9.07E-3	1.51E-5	over
GO:0016168	chlorophyll binding	MF	3.11E-2	7.94E-5	over
GO:0010258	NADH dehydrogenase complex (plastoquinone) assembly	BP	7.53E-5	7.37E-9	over
GO:0044248	cellular catabolic process	BP	1.93E-3	1.13E-6	under
GO:0009768	photosynthesis, light harvesting in photosystem I	BP	2.67E-3	2.07E-6	over
GO:0030026	cellular manganese ion homeostasis	BP	2.67E-3	2.61E-6	over
GO:0042547	cell wall modification involved in multidimensional cell growth	BP	3.47E-3	4.76E-6	over
GO:0030245	cellulose catabolic process	BP	3.47E-3	4.76E-6	over
GO:0071421	manganese ion transmembrane transport	BP	1.06E-2	1.88E-5	over
GO:0044712	single-organism catabolic process	BP	2.54E-2	5.97E-5	under
GO:0010042	response to manganese ion	BP	2.79E-2	6.84E-5	over
GO:0009117	nucleotide metabolic process	BP	4.39E-2	1.20E-4	under

Table S4.3. All identified tef protein sequences mapped to reference canonical biological pathways in KEGG.

Pathway	Sequences in Pathway
Carbon fixation in photosynthetic organisms	174
Glycolysis / Gluconeogenesis	129
Purine metabolism	128
Glyoxylate and dicarboxylate metabolism	128
Pyruvate metabolism	120
Glutathione metabolism	93
Starch and sucrose metabolism	93
Methane metabolism	89
Phenylalanine metabolism	83
Carbon fixation pathways in prokaryotes	82
Glycine, serine and threonine metabolism	80
Phenylpropanoid biosynthesis	80
Alanine, aspartate and glutamate metabolism	79
Cysteine and methionine metabolism	74
Arginine and proline metabolism	70
Amino sugar and nucleotide sugar metabolism	70
Pentose phosphate pathway	65
Oxidative phosphorylation	60
Citrate cycle (TCA cycle)	57
Fructose and mannose metabolism	57
Ascorbate and aldarate metabolism	56
Glycerolipid metabolism	49
Cyanoamino acid metabolism	48
Tryptophan metabolism	47
Glycerophospholipid metabolism	46
Nitrogen metabolism	45
Porphyrin and chlorophyll metabolism	45
Galactose metabolism	40
Carotenoid biosynthesis	37
alpha-Linolenic acid metabolism	36
Metabolism of xenobiotics by cytochrome P450	35
Tyrosine metabolism	35
Propanoate metabolism	35
Pentose and glucuronate interconversions	35
Arachidonic acid metabolism	34
beta-Alanine metabolism	33
Drug metabolism - other enzymes	31
Drug metabolism - cytochrome P450	31
Lysine degradation	31
Pyrimidine metabolism	30
Linoleic acid metabolism	29
Sulfur metabolism	29

N-Glycan biosynthesis	29
Inositol phosphate metabolism	29
Valine, leucine and isoleucine degradation	27
Tropane, piperidine and pyridine alkaloid biosynthesis	27
T cell receptor signalling pathway	27
Taurine and hypotaurine metabolism	26
Ether lipid metabolism	26
Isoquinoline alkaloid biosynthesis	25
Butanoate metabolism	24
Aminobenzoate degradation	22
Ubiquinone and other terpenoid-quinone biosynthesis	21
Fatty acid degradation	21
Phenylalanine, tyrosine and tryptophan biosynthesis	20
Phosphatidylinositol signalling system	20
Histidine metabolism	20
One carbon pool by folate	19
Fatty acid biosynthesis	18
Caprolactam degradation	17
Terpenoid backbone biosynthesis	16
Streptomycin biosynthesis	15
Aminoacyl-tRNA biosynthesis	15
Photosynthesis	14
Sphingolipid metabolism	14
Novobiocin biosynthesis	14
Lysine biosynthesis	14
Selenocompound metabolism	13
Other glycan degradation	11
Glycosphingolipid biosynthesis - globo series	11
Valine, leucine and isoleucine biosynthesis	10
Monoterpenoid biosynthesis	10
Synthesis and degradation of ketone bodies	10
Benzoate degradation	10
Vitamin B6 metabolism	10
Flavonoid biosynthesis	10
mTOR signalling pathway	10
Toluene degradation	9
Xylene degradation	9
Glycosphingolipid biosynthesis - ganglio series	9
Biosynthesis of terpenoids and steroids	9
Glycosaminoglycan degradation	9
Thiamine metabolism	8
Chloroalkane and chloroalkene degradation	7
Pantothenate and CoA biosynthesis	7
Biosynthesis of ansamycins	6

Retinol metabolism	6
Biosynthesis of unsaturated fatty acids	5
Carbapenem biosynthesis	5
Various types of N-glycan biosynthesis	5
Limonene and pinene degradation	5
Flavone and flavonol biosynthesis	5
Aflatoxin biosynthesis	5
Tetracycline biosynthesis	5
Riboflavin metabolism	5
Biotin metabolism	4
Isoflavonoid biosynthesis	4
Steroid hormone biosynthesis	4
Polyketide sugar unit biosynthesis	3
C5-Branched dibasic acid metabolism	3
Biosynthesis of vancomycin group antibiotics	3
Geraniol degradation	2
Cutin, suberine and wax biosynthesis	2
Glycosylphosphatidylinositol(GPI)-anchor biosynthesis	2
Fatty acid elongation	2
Ethylbenzene degradation	2
Steroid degradation	1
Caffeine metabolism	1
Lipoic acid metabolism	1
Butirosin and neomycin biosynthesis	1
Sesquiterpenoid and triterpenoid biosynthesis	1
Zeatin biosynthesis	1
Diterpenoid biosynthesis	1
Indole alkaloid biosynthesis	1
Glucosinolate biosynthesis	1
Nicotinate and nicotinamide metabolism	1
Benzoxazinoid biosynthesis	1
Steroid biosynthesis	1
Styrene degradation	1
Folate biosynthesis	1
Glycosaminoglycan biosynthesis - heparan sulfate / heparin	1
Total number of protein sequences mapped:	**3438**

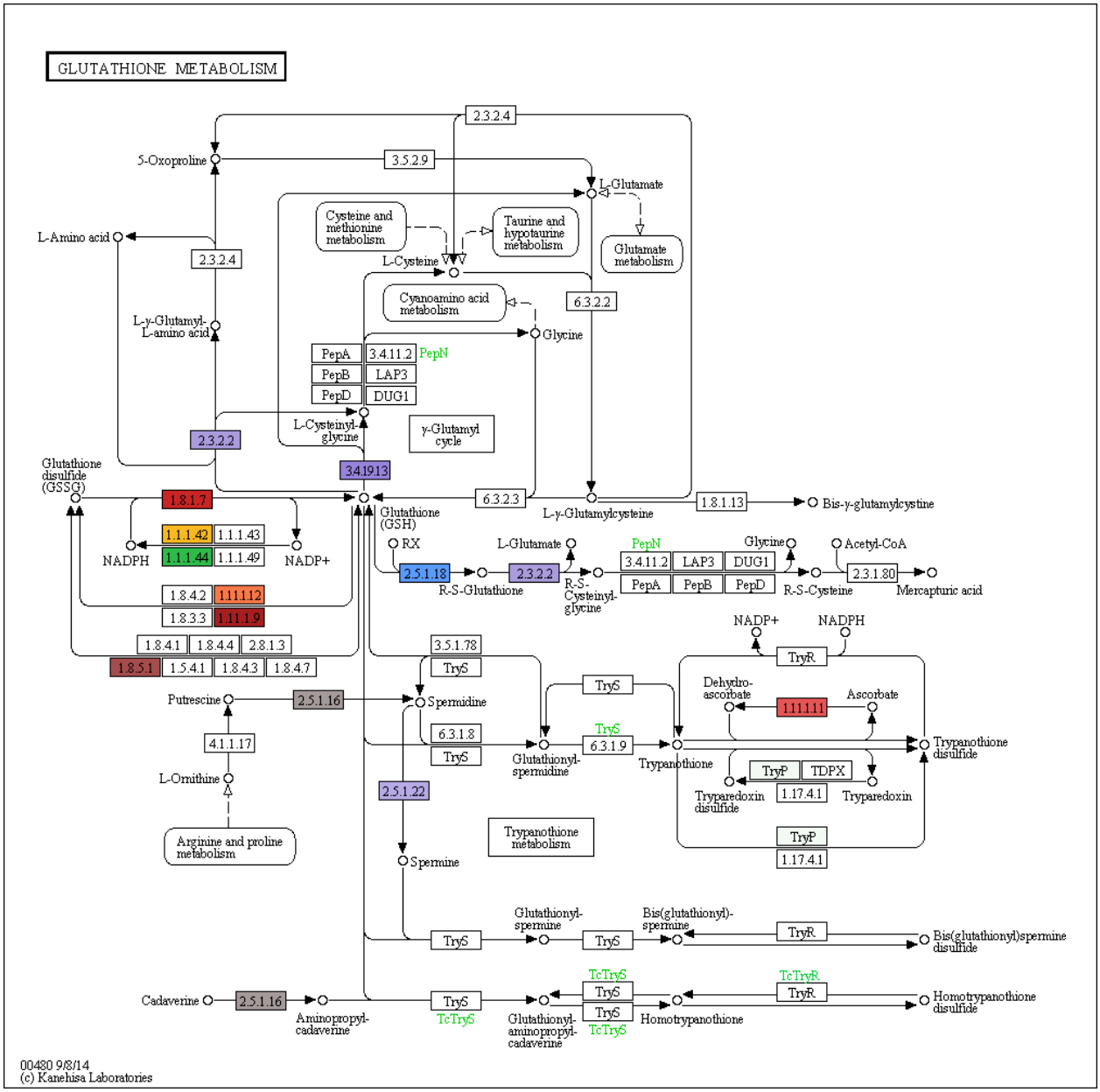

Fig. S4.1 Glutathione metabolism pathway involving tef protein sequences retrieved from KEGG.

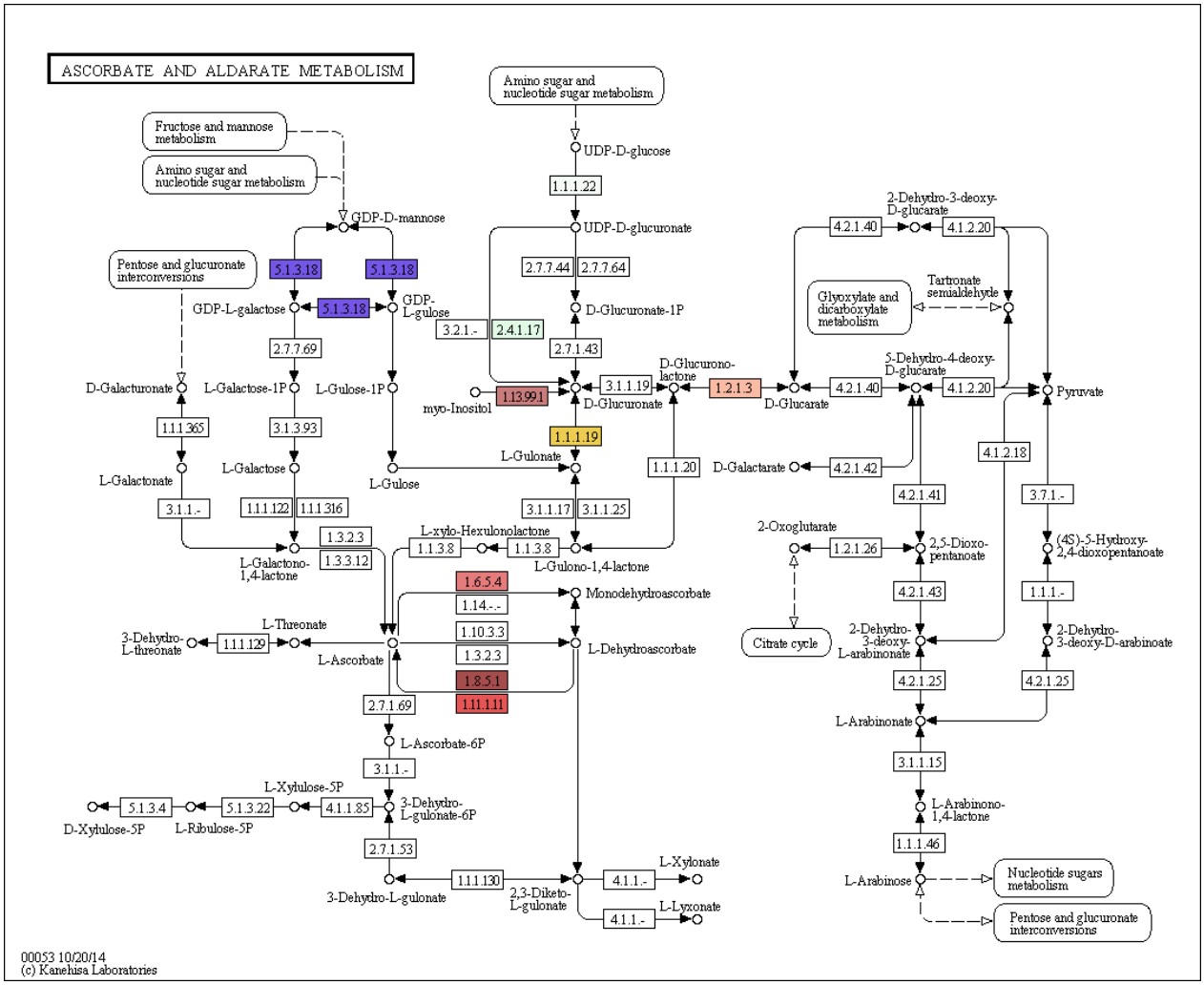

Fig. S4.2 Ascorbate and aldarate metabolism pathway involving tef protein sequences retrieved from KEGG.

Chapter 5: Biological validation of tef proteins

5.1 Introduction

Although protein identification and quantification by iTRAQ analysis is a well-established and highly sensitive tool (Agrawal *et al.*, 2013; Liu *et al.*, 2014), sample variance, outliers and false positive protein identification is a major concern (Gan *et al.*, 2007). In addition, due to the thousands of spectra generated through mass spectrometry approaches, most experiments require stringent manual validation procedures (Gan *et al.*, 2007). However, because this is a costly and time consuming process, it is more practical to manually validate a sub-set of the data acquired and infer biological confirmation from those results. To achieve this and to investigate the biological response of tef protein accumulation during dehydration stress, two methods of biological validation were performed, *viz.* western blots for verification of protein presence and changes therein as well as relevant physiological assays to determine activity of proteins selected at the various water contents. The biological validation tests were executed on various high-abundance proteins generated from the three statistically analysed (foreground) protein datasets in Chapter 3, namely Tef Extended (TE), Tef Extended-unique (TEU) and Monocot-unique (MU) in order to have a well-represented biological indication of tef proteomic dehydration stress response.

For western blot biological validation, two proteins were chosen from the Tef Extended (TE, Table 3.1), Tef Extended-unique (TEU, Table 3.3) and Monocot-unique (MU, Table S3.1) lists of high-abundance proteins. These included: fructose-bisphosphate aldolase (FBA), present in the TE and TEU datasets; glutamine synthetase (GLN), present in the MU dataset and superoxide dismutase (SOD), not present in any of the differentially regulated foreground protein lists but a ROS-scavenging enzyme that is commonly increased in abundance under water-deficit conditions (Farrant *et al.*, 2007) and for which the antibody was available for use. In addition, the enzymatic activities of high-abundance proteins were observed throughout dehydration stress. These included: fructose-bisphosphate aldolase (FBA) and monodehydroascorbate reductase (MDHAR), present in the TE and TEU datasets; peroxidase 3 (POX), present in the TE dataset and glutamine synthetase (GLN), present in the MU dataset.

5.2 Materials and Methods

5.2.1 Plant material

Seven-week-old tef plants were dehydrated as described in Chapter 2 (section 2.2.2.1, Fig. 2.1), during which leaves were sampled from fully hydrated and variously dehydrated plants for the assays described in this chapter. For testing of western blot analyses, RWC points similar to those used in iTRAQ analysis (section 3.2.1, Fig. 3.1) were chosen, *viz.* 92% RWC (hydrated-control) and 55, 52 and 50% RWC (dehydrated-experimental repeats designated D1, D2 and D3, respectively) from different biological repeats of pooled plants. For testing of enzyme activities, leaf tissues in the RWC ranges: 90-95, 75-80, 60-65, 50-55, 35-40 and 25-30% RWC, were selected from different biological repeats of pooled plants and assayed using spectrophotometric methods (described below).

5.2.2 Western blot analyses

The three proteins chosen for immunodetection were subjected to PAGE separation and subsequent western blotting with use of commercial antibodies (Agrisera; As08 170, As08 294 and As08 295).

Total proteins were extracted and quantified as described in section 3.2.2 and 3.2.3, respectively. Thereafter, protein extracts in 2% (w/v) SDS were mixed with 2x Laemmli sample loading buffer (Sigma-Aldrich, Inc.) and boiled at 90 °C for 5 min to facilitate protein denaturation. The re-suspended protein sample was then loaded onto 12% sodium dodecyl sulphate polyacrylamide gels (30% (v/v) acrylamide, 0.375M Tris-HCl, pH 8.8, 0.1% (w/v) SDS, 0.05% (w/v) ammonium persulfate and TEMED) at a concentration of 15 µg. Separation of proteins and molecular weight marker (Fermentas, USA) by SDS-PAGE was performed at room temperature in the Mini-Protean Tetra Cell (Bio-Rad Laboratories, Inc.) at constant voltage of 100 V for 2 h in SDS running buffer (25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS). Once electrophoresis was complete the gels were carefully removed from glass plates and either stained for quality assessment or prepared for western blotting.

Gels were stained in Coomassie Brilliant Blue (CBB) R-250 staining solution (45% (v/v) methanol, 10% (v/v) acetic acid and 0.02% (w/v) Coomassie Brilliant Blue R-250) on a shaker at 37 °C for 1 h, followed by destaining in destain solution (40% (v/v) methanol and 10% (v/v) acetic acid) overnight until proteins bands and molecular weight marker were clearly visible. For western blotting analysis, gels were incubated in cold Tris buffered saline (TBS) (50 mM Tris-HCl, pH 7.5, 150 mM NaCl) at 4 °C for approximately 5 min to remove residual SDS. Gels were then transferred to pure nitrocellulose membrane (PALL Life Sciences, USA) pre-soaked in TBS together with Whatmann 3 MM filter paper and sandwiched tightly together with transfer sponges in transfer cassette. Transfer of proteins to membrane was conducted using the Criterion™ Blotter apparatus (Bio-Rad Laboratories, Inc.) at 100 V for 1 h at 4 °C.

To confirm efficient transfer of proteins to membrane, membranes were stained in Ponceau S reversible total protein stain (0.1% (w/v) Ponceau S powder, 5% (v/v) acetic acid) on a shaker for 1 min, until protein lanes were visible. Excess Ponceau S stain was removed by rinsing with distilled water until membranes were clear. To observe equal loading of proteins in lanes and to capture the loaded molecular weight marker, the Bio-Rad ChemiDoc™ XRS imager with 'colorimetric' settings, was used. The membranes were then placed in blocking buffer (5% (w/v) fat-free milk powder in TBS containing 1% (v/v) Tween-20) for 1 h at 4 °C to block non-specific proteins. After blocking, membranes were incubated in respective primary antibodies diluted in blocking buffer at the concentrations displayed below (Table 5.1) for 16 h at 4 °C, with mild agitation.

Table 5.1 The detection parameters including primary and secondary antibody dilutions for each antibody used in western blot analysis of selected tef protein targets in response to dehydration stress.

Protein target	**Protein concentration (µg)**	**Primary antibody dilution**	**Secondary antibody dilution**
SOD (Cu/Zn-SOD, chloroplastic)	15	1:1500	1:5000
FBA	15	1:5000	1:5000
GLN (GLN1-cytoslic+GLN2-chloroplastic)	15	1:10 000	1:5000

To remove all traces of unbound primary antibody, membranes were washed a total of four times in TBS-T (Tris buffered saline containing 1% (v/v) Tween-20) at 5 min intervals each time with mild agitation at 4 °C, before being incubated with goat anti-rabbit peroxidase conjugated secondary antibody (Agrisera, AS09 602) in blocking buffer at the concentrations stated above (Table 5.1) for 1 h at 4 °C, with mild agitation. Following secondary antibody incubation, membranes were once again washed at least four times with TBS-T for 5 min intervals with mild agitation at 4 °C to remove all traces of unbound secondary antibody.

Thereafter, detection and visualisation of protein expression was performed using the WesternBright ECL HRP chemiluminescent detection kit (Advansta, USA), following the manufacturer's instructions for solutions and applying the mixed substrate directly to membranes. The membrane images were visualised by chemiluminescence using the ChemiDoc™ XRS imager installed with ImageLab software (version 4.1). For relative quantification of detected band intensities, the 'Volume analysis' tool in ImageLab was used to create a rectangular area of the same size for each band intensity present, using a global background subtraction of the whole membrane. These relative quantification values, corrected for background subtraction and made relative to the hydrated-control (92% RWC) band intensity, were used for statistical analysis and graph generation using GraphPad Prism 6.0 software. A one-way ANOVA statistical test was performed in GraphPad Prism with the relative quantification values and statistical significance of higher or lower intensity bands were established according to p-value (p-value < 0.05). In addition, at least five biological repeats for each antibody were conducted as stated above.

5.2.3 Enzyme assays

The enzyme activities of monodehydroascorbate reductase (MDHAR, EC: 1.6.5.4), fructose-bisphosphate aldolase (FBA, EC: 4.1.2.13), peroxidase (POX, EC: 1.11.1.7) and glutamine synthetase (GLN, EC: 6.3.1.2) were assayed using spectrophotometric methods as described below. GraphPad Prism 6.0 software was used for graph generation and statistical analysis by one-way ANOVA, where significance was based on p-value (p-value < 0.05) of changing enzyme activities at differing RWC ranges.

Total protein contents in extracts for all assays performed, were determined using Bradford's method of protein quantification with BSA as a standard (Quick Start Bradford Protein Assay), following the manufacturer's instructions for microplate standard assay procedures. Total protein absorbance was measured at 595 nm using the MultiSkan EX microplate reader (Thermo Fisher Scientific, Inc.,USA) and total enzyme activities for assays were measured using the cuvette and microplate reader MultiSkan GO, (Thermo Fisher Scientific, Inc., USA).

5.2.3.1 Monodehydroascorbate reductase (MDHAR, EC: 1.6.5.4)

Enzyme extraction was performed according to Valyova *et al.* (2012) with modifications. Approximately 0.25 g leaf tissue was ground in liquid nitrogen to a fine powder and 0.2% (w/w) insoluble PVPP was added. A volume of 2.5 ml extraction buffer (0.05 M KH_2PO_4 buffer, pH 7.0, 1 mM Ascorbate, 1 mM EDTA) was added to ground material and vortexed for 5 min before centrifugation at 12,000 *x g* at 4 °C. The supernatant was passed through a PD-10 de-salting column (GE Healthcare, USA) previously equilibrated with extraction buffer according to manufacturer's instructions. Once the supernatants were passed through the column, purified extracts were eluted with 3 ml extraction buffer and quantified for total protein concentrations as stated above (section 5.2.3).

MDHAR enzyme activity was determined as originally described by Miyake and Asada (1992) and further employed by Kingston-Smith and Foyer (2000). The decrease in absorbance at 340 nm due to the oxidation of NADH was observed in a reaction mixture of 1 ml containing (50 mM Hepes-KOH buffer, pH 7.3, 0.1 mM NADH, 2.5 mM Ascorbate, 100 µl extract). Once sample extract was added, the spectrophotometer was zeroed and sample absorbance (without enzyme) was measured every 10 sec over a 3 min interval to provide a blank for enzyme activity. Thereafter, 0.45 units of ascorbate oxidase enzyme (EC: 1.10.3.3, Sigma-Aldrich, Inc.) in 100 µl distilled water was added to begin the reaction which was then left to proceed for 3 min at 25 °C, again with measurements every 10 sec. The rate of enzyme activity was then determined by subtracting the blank absorbance from sample absorbance and calculating the change in absorbance over time to obtain a rate. Enzyme activity was calculated using the rate of change with the following formula:

$$\text{Total enzyme activity} = \frac{(\text{rate, abs.s}^{-1}) \text{ X (final volume in cuvette, 1 ml)}}{(\text{extract volume, 0.1 ml}) \text{ X (extinction coefficient of NADH, 6.22 mM}^{-1}\text{cm}^{-1})}$$

Specific activity is given as enzyme units.mg protein^{-1}.

5.2.3.2 Fructose-bisphosphate aldolase (FBA, EC: 4.1.2.13)

Extraction and analysis of fructose-bisphosphate aldolase enzyme activity was performed according to Mundree *et al.* (2000) with modifications. Approximately 0.25 g leaf tissue was ground in liquid nitrogen to a fine powder and 0.2% (w/w) insoluble PVPP was added. A volume of 2.5 ml extraction buffer (0.05M KH_2PO_4 buffer, pH 7.0, 4 mM $MgCl_2$, 1 mM EDTA, 10% (v/v) glycerol, 5 mM dithiothreitol (DTT)) was added to ground material and vortexed for 5 min before centrifugation at 12,000 *x g* at 4 °C. The resulting supernatant was then passed through a PD-10 de-salting column as detailed in section 5.2.3.1 and eluents quantified for total protein concentration as described in section 5.2.3.

FBA enzyme activity was measured in a combined reaction with glycerol-3-phosphate dehydrogenase (G-3-P) (EC: 1.1.1.8, Sigma-Aldrich, Inc.) and triose-phosphate-isomerase (T-P-I) (EC: 5.3.1.1, Sigma-Aldrich, Inc.) in the forward reaction at 22 °C by observing the decrease in absorbance at 340 nm due to the oxidation of NADH. A reaction mixture of 1 ml containing (50 mM Hepes-KOH buffer, pH 7.3, 0.1 mM NADH, 1 mM EDTA, 0.75 units G-3-P and 10 units T-P-I diluted in 250 µl dH_2O) was reconstituted and sample extract (100 µl) was added. Once extract was added to reaction mixture, the spectrophotometer was zeroed and samples (without substrate) were read at 340 nm by taking a measurement every 1 min for 12 min. To start the reaction, 4 mM fructose-bisphosphate (Sigma-Aldrich, Inc.) substrate was added and the reaction was monitored every 1 min for 12 min at 340 nm. As for MDHAR, the rate of change in enzyme activity was calculated by firstly subtracting the blank absorbance from sample absorbance at each time point and subsequently calculating the rate of change. The rate of change in enzyme activity was then used in the equation above (section 5.2.3.1) for calculating total enzyme activity using the extinction coefficient of NADH (6.22 $mM^{-1}cm^{-1}$) in a 1 cm^3 cuvette. One unit of activity was defined as the amount of enzyme required for the oxidation of 2 µmol NADH at 22 °C and specific activity was given by enzyme units.mg $protein^{-1}$.

5.2.3.3 Peroxidase (POX, EC: 1.11.1.7)

Extraction and assay of total peroxidase activity was performed according to Dionisio-Sese and Tobita (1998), by the determining the rate of guaiacol oxidation (Chance and Maehly, 1955). Approximately 0.15 g leaf tissue was ground in liquid nitrogen to a fine powder and 1.5 ml cold extraction buffer (0.1 M KH_2PO_4 buffer, pH 6.0) was added and mixed by vortexing for 5 min before centrifugation at 12,000 *x g* at 4 °C. The resulting supernatant (crude extract) was kept cold at all times and quantified according to section 5.2.3, before being immediately used in enzyme assays. The reaction mixture containing (10 mM KH_2PO_4 buffer pH 6.0, 8 mM guaiacol and 100 µl extract) had a final volume of 3 ml. To start the reaction, 2.75 mM hydrogen peroxide (H_2O_2, Saarchem) was added and the increase in absorbance due to guaiacol oxidation was measured at 470 nm every 20 sec over a time period of 3 min. The change in absorbance at 470 nm per min (rate of activity) was used in enzyme activity calculation. Total enzyme activity was calculated using the equation displayed in section 5.2.3.1, with the extinction coefficient of tetraguaiacol (26.6 $mM^{-1}cm^{-1}$) and adjusting the final volume in the cuvette to 3 ml. A unit of peroxidase

activity was expressed as the amount of enzyme required to catalyse the conversion of 1 mmol H_2O_2, with guaiacol as hydrogen donor, per min under specified conditions, while specific activity was given as enzyme units.mg protein^{-1}.

5.2.3.4 Glutamine synthetase (GLN, EC: 6.3.1.2)

Glutamine synthetase extraction and enzyme activity was determined according to the method by Rhodes *et al.* (1975) and further employed by Machado *et al.* (2001). Approximately 0.25 g leaf tissue was ground in liquid nitrogen to a fine powder adding 0.2% (w/w) insoluble PVPP. A volume of 2.5 ml cold extraction buffer (0.1 M imidazole-HCl buffer, pH 7.8, containing 1 mM DTT) was added to ground material and vortexed for 5 min before centrifugation at 12,000 *x g* at 4 °C. The resulting supernatant was then passed through a PD-10 de-salting column as explained in section 5.2.3.1 and eluents quantified for total protein concentrations as explained in section 5.2.3.

GLN enzyme activity was measured through the forward (synthetase) reaction by observing the formation of the product ɣ-glutamyl hydroxamate, from the reaction of glutamate and hydroxylamine with ATP at 535 nm. The reaction mixture in a total volume of 250 µl consisting of (0.05 M imidazole buffer, pH 7.4, 0.5 M monosodium glutamate (MSG), 0.06 M hydroxylamine, 0.2 M $MgSO_4$ and 100 µl extract) was started by the addition of 0.024 M ATP (Sigma-Aldrich, Inc.). Thereafter samples were incubated in a water bath at 32 °C for 30 min to allow the reaction to proceed. To stop the reaction, Ferguson and Sims (1971) stop reagent (0.67 M HCl, 0.2 M TCA and 0.37 M $FeCl_3$) at an equivalent volume to that of reaction mixture (250 µl), was added and the resulting precipitate was removed by centrifugation at 3,000 *x g* for 5 min. For the blank (control) sample, the same procedure was followed, however, the substrate, MSG was omitted and replaced with an appropriate volume of 0.05 M imidazole buffer, pH 7.4. All supernatants retrieved after centrifugation were measured at 535 nm using a microplate reader with the control sample (no substrate) as a blank. GLN concentrations were determined through the use of a standard curve with ɣ-glutamyl hydroxamate (Sigma-Aldrich, Inc.) as a standard. The standards were prepared in the same manner as samples, but replacing sample extract with the product, ɣ-glutamyl hydroxamate at known concentrations in a concentration range of 1.25 to 20 mM ɣ-glutamyl hydroxamate. One unit of GLN activity was defined as the amount of enzyme required to produce 1 mmol ɣ-glutamyl hydroxamate product per min at specified conditions and specific activity was given as mM product formed.min^{-1}.mg protein^{-1}.

5.3 Results and Discussion

5.3.1 Western blot analyses

To test the accumulation of the chosen high-abundance proteins in response to dehydration stress, western blotting was used as one of the methods of biological validation of data generated from iTRAQ analysis (in Chapter 3). The protein targets chosen to observe the change in band intensities at designated molecular weights (kDa) for the chosen antibodies in response to dehydration stress are shown in Figure 5.1 below.

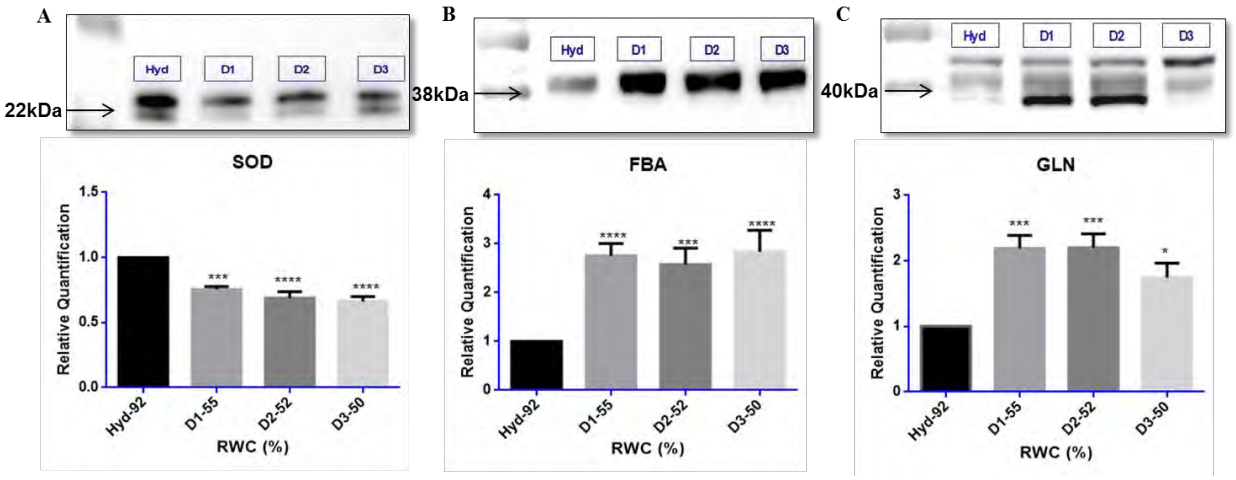

Fig. 5.1 Western blot biological validation of chosen high-abundance proteins: superoxide dismutase–SOD (A), fructose bisphosphate aldolase–FBA (B) and glutamine synthetase–GLN (C). Western blots shown are at respective band sizes for all antibodies tested, for both control (Hyd-92% RWC) and dehydrated repeats (D1-55, D2-52 and D3-50% RWC). Relative quantification of band intensities for more than 5 western blots ($n \geq 5$) were performed using ImageLab software (Bio-Rad Laboratories, Inc.) and analysed for statistical significance through one-way ANOVA (p-value < 0.05), shown by asterisks placed on columns (RWC points) statistically significant to control (Hyd-92% RWC). Error bars represent standard error.

SOD (Cu/Zn-SOD) (22 kDa, Fig. 5.1A), was shown to be significantly decrease in protein abundance in response to dehydration stress (p-value < 0.05). A decrease in band intensities and relative quantification values for dehydrated repeats (D1-55, D2-52 and D3-50% RWC) were displayed in comparison to control (Hyd-92% RWC). In contrast, the protein targets FBA and GLN, were shown to be significantly increase in protein abundance in response to dehydration stress (p-value < 0.05) (Fig. 5.1B, C). These proteins displayed enhanced band intensities and significant increases in relative quantification values, by exhibiting at least a 2-fold increase at 38 kDa for FBA (Fig. 5.1B) and 40 kDa for GLN (Fig. 5.1C), when comparing dehydrated repeats (D1-55, D2-52 and D3-50% RWC) to control (Hyd-92% RWC).

Although SOD appears to be decrease in abundance in tef in response to dehydration stress (Fig. 5.1A), its activity is known to be commonly up-regulated under water-deficit conditions (Farrant *et al.*, 2007). The SOD group of metalloenzymes functions in early ROS detoxification, by catalysing the dismutation of superoxide (O_2^-), one of the first ROS to be produced, into oxygen and hydrogen peroxide (H_2O_2) in response to stress conditions (McCord and Fridovich, 1969; Bowler *et al.*, 1992; Cruz de Carvalho,

2008). However, since this reaction, essentially only converts from one form of ROS to another (O_2^- to H_2O_2), H_2O_2 also needs to be detoxified as its presence and accumulation attacks thiol proteins (Noctor *et al.*, 1998; Cruz de Carvalho, 2008). The antioxidant enzymes, catalase, ascorbate peroxidase and glutathione reductase (part of the ascorbate-glutathione cycle) then function together in a signalling network to further detoxify H_2O_2 accumulation to H_2O and O_2 and thereby scavenge excessive ROS formation during stress conditions (Bowler *et al.*, 1992; Foyer *et al.*, 1994; Qureshi *et al.*, 2007; Cruz de Carvalho, 2008). It is possible, that early induction of SOD protein had occurred during the early stages of dehydration in tef, at water contents above 55% RWC; however, since western blotting was tested at the dehydrated water contents of 50-55% RWC only, increased SOD protein expression could have been missed.

Alternatively, the regulation of SOD protein has been reported in numerous crop plant studies where various forms of SOD, based on organelle localisation, are either up or down-regulated in response to dehydration stress (Salekdeh *et al.*, 2002; Hajheidari *et al.*, 2005; Qureshi *et al.*, 2007). In this instance, the metal co-factor (Cu/Zn) form of SOD, localised in the chloroplast organelle was detected and the cytoplasmic form of Cu/Zn-SOD was not investigated during western blotting. In a study with sugar beet (*Beta vulgaris* L.) leaves under dehydration stress conditions, the cytosolic form of Cu/Zn-SOD was significantly up-regulated (Hajheidari *et al.*, 2005). While in a study observing root tissue exposure of Indian mustard (*Brassica juncea* L.) to the heavy metal, cadmium, Cu/Zn-SOD was found to be down-regulated (Alvarez *et al.*, 2009). In rice, however, a contrasting result was observed, where Cu/Zn-SOD in the chloroplast tissues were down-regulated and the cytosolic form of Cu/Zn-SOD was up-regulated in response to dehydration stress (Salekdeh *et al.*, 2002). This suggests that, depending on the tissue and organelle under study, SOD could respond differently in response to dehydration stress.
Further studies have suggested that Cu/Zn-SOD down-regulation and its role in signalling is linked to the regulation of lignification in soybean roots (*Glycine max* L.) under flooding stress (Komatsu *et al.*, 2010; Nanjo *et al.*, 2011). Komatsu *et al.* (2010) suggested that soybean roots and hypocotyls respond through signalling cascades triggered by the accumulation of H_2O_2 and ascorbate that cause a decline in soybean lignification through a reduction in polysaccharide linkages (Nanjo *et al.*, 2011). Thus, the regulation of SOD and other ROS scavenging enzymes could potentially be associated with adaptive processes in crop plants in response to various abiotic stress factors (Nanjo *et al.*, 2010; 2011).

The protein targets FBA and GLN, key enzymes involved in carbohydrate and nitrogen metabolism, respectively, were significantly increased in protein abundance during western blotting in response to dehydration stress (Fig. 5.1B, C) supporting our iTRAQ findings (in Table 3.1, Table 3.3 and Table S3.1, Chapter 3). A triplet band is displayed for GLN, where total GLN activity, GLN 1 (cytosolic at 39-40 kDa) and GLN2 (chloroplastic at 44-45 kDa) were detected (Fig. 5.1C). In the dehydrated repeats D1 and D2 (55 and 52% RWC, respectively), the cytosolic form of GLN at approximately 40 kDa (GLN1) was

enhanced, while in D3 (50% RWC), the chloroplastic form (GLN2) was more prevalent. Although different forms of GLN are depicted, perhaps representing different isoforms of GLN, the overall trend of increased protein abundance and enhanced band intensity is observed with dehydration stress, coinciding with the iTRAQ findings in tef (Table S3.1, Chapter 3).The accumulated proteins, FBA and GLN, enhanced with imposed dehydration stress were then further investigated by means of physiological enzyme assays together with two additional stress responsive enzymes shown to increase in protein abundance in response to dehydration stress (section 5.3.2).

5.3.2 Enzyme assays

A total of four enzymes, namely MDHAR, FBA, POX and GLN, were tested for their activity in tef plants during dehydration stress (Fig. 5.2). Of the enzymes tested, three were shown to significantly increase in activity during dehydration stress, namely MDHAR, FBA and POX at 60-65% RWC (Fig. 5.2A to C), while GLN showed a decrease in enzyme activity (Fig. 5.2D). Low levels of GLN activity were maintained throughout dehydration (60-65% RWC to 25-30% RWC) (Fig. 5.2D).

The stress responsive antioxidant enzymes known to offer protection against free radical accumulation, MDHAR and POX (Zhang and Kirkham, 1994; Shah *et al.*, 2001; Gill and Tuteja, 2010; Huang *et al.*, 2013; Shin *et al.*, 2013), displayed a large increase in enzymatic activity at 60-65% RWC, where starting concentrations of enzymes increased approximately two-fold (Fig. 5.2A, C). A second spike or increase in enzyme activity was observed for both MDHAR and POX at 35-40 and 25-30% RWC, respectively, towards the latter stages of dehydration (Fig. 5.2A, C).

MDHAR was significantly increased in quantitative expression during iTRAQ analysis in response to dehydration stress in the TE (Table 3.1, p-value = 0.006) and TEU datasets (Table 3.3, p-value = 0.027, Chapter 3) and shown to be largely enriched in the GO-term analysis of high-abundance proteins (Table 4.2, A; FDR = 1.40E-3, p-value = 1.37E-7), when conducting bioinformatics evaluation (Chapter 4). Furthermore, MDHAR was further highlighted during KEGG pathway analysis when the ascorbate and aldarate metabolism pathway was investigated (Table 4.3, B, EC:1.6.5.4; Fig. 4.6B, Chapter 4).

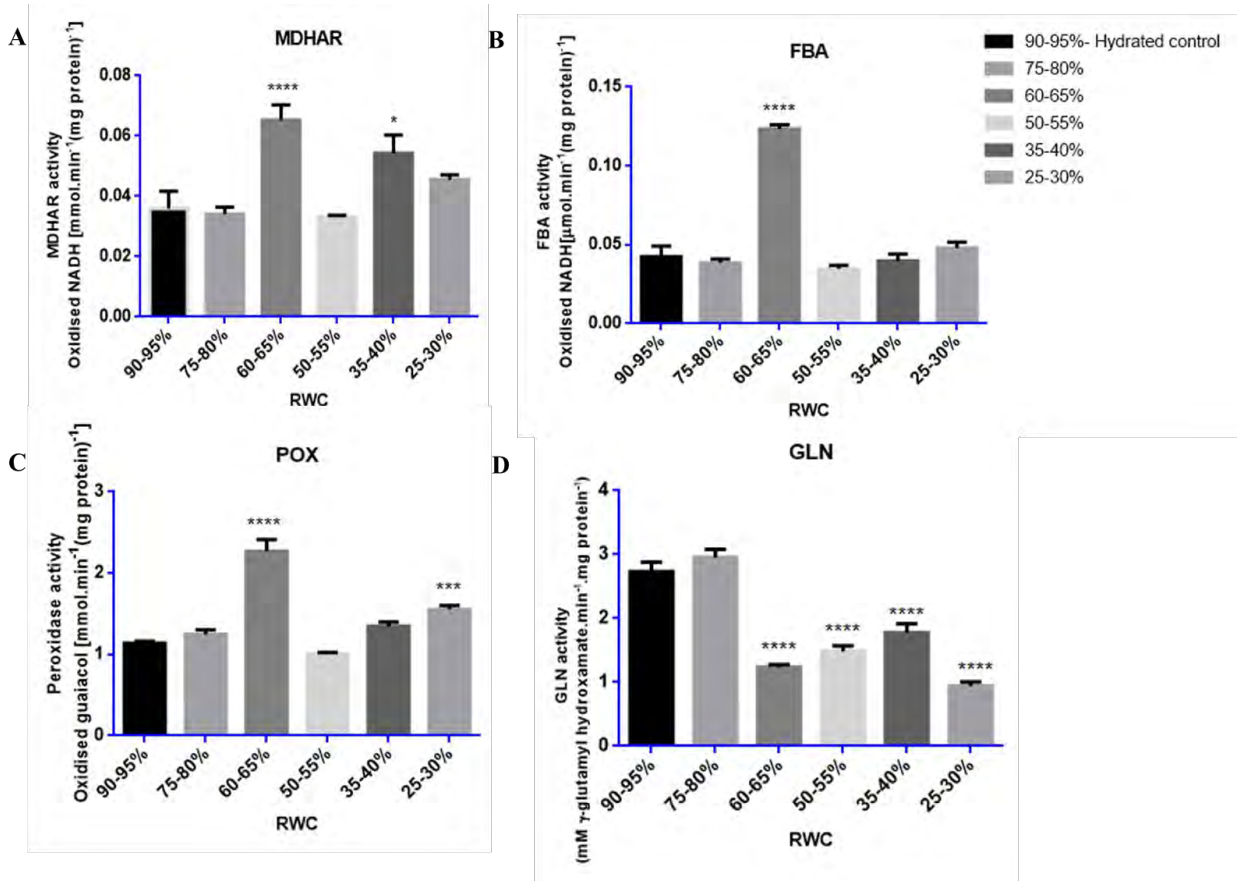

Fig. 5.2 Enzyme assays of selected high-abundance proteins: monodehydroascorbate reductase-MDHAR (A), fructose bisphosphate aldolase-FBA (B), peroxidase–POX (C) and glutamine synthetase–GLN (D). Enzyme activities are displayed as specific activity (enzyme units.mg protein^{-1}) and were measured in tef leaves throughout dehydration stress (90-95% to 25-30% RWC). Enzyme assays were performed on more than 10 replicates ($n \geq 10$) from each RWC range. MDHAR and FBA activities were measured by monitoring the rate of NADH oxidation at 340 nm (A and B , respectively), while POX activity was measured by following the rate of guaiacol oxidation at 470 nm (C) and GLN activity was measured through the synthetase (forward) reaction by observing the amount of product (ɣ-glutamyl hydroxamate) formed.min^{-1}, at 535 nm (D). Statistical significance (one-way ANOVA, p-value < 0.05) is shown by asterisks placed on columns (RWC ranges) statistically significant to hydrated-control (90-95% RWC). Error bars denote standard error between tested replicates.

It would seem that the generation of ascorbate as well as the regulation and maintenance of the ascorbate-glutathione cycle are of empirical importance in tef in response to dehydration stress. Previous studies have shown that if the appropriate concentrations of ascorbate and glutathione are not maintained, particularly the ratios between glutathione disulphide/glutathione (GSSG/2GSH) and the dehydroascorbate/ascorbate redox couples (Noctor *et al.*, 1998), then a signalling cascade is triggered that results in programmed cell death (PCD) (Kranner *et al.*, 2006; Kamies *et al.*, 2010). Therefore, it could be hypothesised that necessary concentrations of ascorbate and glutathione in tef need to be maintained to avoid this mechanism. However, further studies regarding PCD and its relation to all the enzymes in the ascorbate-glutathione cycle (ascorbate peroxidase, monodehydroascorbate reductase, dehydroascorbate reductase and glutathione reductase) would need to be pursued.

FBA catalyses the reversible conversion of glyceraldehyde-3-phosphate and dihydroxyacetone phosphate to fructose-1, 6-bisphosphate during glycolysis/gluconeogenesis or in the reaction where erythrose-4-phosphate and dihydroxyacetone phosphate is converted to sedoheptulose-1,7-bisphosphate in the Calvin cycle (Uematsu *et al.*, 2012). In addition, FBA is one of the six non-regulated enzymes in the Calvin cycle, where their activity is based on expressional regulation and protein degradation rather than post-translational modifications (Graciet *et al.*, 2004; Uematsu *et al.*, 2012). A significant increase in FBA activity was observed in tef at 60-65% RWC, where enzyme activities were shown to increase more than two-fold (Fig. 5.2B). Subsequent to high levels of FBA at the 60-65% RWC, activity was shown to return to starting levels and was maintained throughout dehydration.

The increase in FBA activity has been observed in stress response for various other crop plants such as rice, in response to drought stress and increased salinity (Salekdeh *et al.*, 2002; Salekdeh and Komatsu, 2007); wheat seedlings, in response to anaerobic conditions (Kamal *et al.*, 2012; Komatsu *et al.*, 2014); wheat roots, in response to increased aluminium concentrations (Oh *et al.*, 2014; Komatsu *et al.*, 2014) and in Indian mustard, in response to increase cadmium concentrations (Alvarez *et al.*, 2009). Previous investigations of FBA in the resurrection plant, *Xerophyta visosa* Baker, subjected to dehydration stress showed an increase in FBA activity as RWC values of leaves decreased (Mundree *et al.*, 2000). FBA activity was maintained to a low water content of 5% RWC and was suggested to play an active metabolic role in tolerance to dehydration stress in *X. viscosa* (Mundree *et al.*, 2000). The anabolic and catabolic use of soluble sugars perhaps through the expressional regulation of FBA and similar enzymes active in carbohydrate metabolism, can be proposed to have a multiplexed, metabolic role in tef plant maintenance during dehydration stress.

An increase in POX activity has been related to many oxidative and abiotic stresses (Sreenivasulu *et al.*, 1999; Mittler *et al.*, 2004) particularly in response to dehydration stress conditions in the crop plants, wheat (Zhang and Kirkham, 1994; Chakraborty and Pradhan, 2012), oilseed rape (Abedi and Pakniyat, 2010), sunflower (Nazarli *et al.*, 2011), horse gram beans (Murthy *et al.*, 2012) as well as in response to salt stress in fox-tail millet and rice (Dionisio-Sese and Tobita, 1998; Sreenivasulu *et al.*, 1999). The increased production of free radicals as a consequence of stress conditions has been proposed to be the main reason for membrane lipid peroxidation, whereby the extent of peroxidation-induced damage is regulated by the antioxidative peroxidase enzyme system (Sreenivasulu *et al.*, 1999; Shah *et al.*, 2001). This could be due in part to the ability of POX acting on increased levels of H_2O_2 in cells as dehydration stress proceeds, even towards the final stages of dehydration stress (25-30% RWC) (Fig. 5.2C). The free radical, H_2O_2, has been postulated to have a dual role in plant cells, by either acting as a signalling molecule at low concentrations during non-stress conditions or as an activator of PCD at high concentrations during stressed conditions (Quan *et al.*, 2008; Gill and Tuteja, 2010). A clear indication of dehydration stress-induced injury, in the form of membrane damage was observed in tef by both

electrolyte leakage measurements (Fig. 2.2) and ultra-structural investigations (Fig. 2.5B to D; Fig 2.6A to C), particularly towards the latter stages of dehydration. This is perhaps due to the extenuating effects of H_2O_2 build up.

Of the enzymes tested, GLN was the only enzyme shown to decrease in activity with imposed dehydration stress in tef (Fig. 5.2D). GLN plays a pivotal role in the assimilation of nitrogen in the form of ammonium into amino acids and various other reduced nitrogenous compounds in plants (Miflin and Habash, 2002; Molina-Rueda *et al.*, 2013). GLN has been likened to the enzyme Rubisco in carbohydrate metabolism, by having the same important functionality in the assimilation of nitrogen (Teixeira and Pereira, 2007; Nagy *et al.*, 2013). Due to the uptake of nitrogen in plants being greatly influenced by soil water availability (Quaye *et al.*, 2009; Molina-Rueda *et al.*, 2013), the influence of low soil water contents as a consequence of drought conditions, would have a largely negative impact on nitrogen assimilation in plant tissues (Molina-Rueda *et al.*, 2013).

A significant number of crop plants studied have been reported to show a decrease in GLN activity, particularly in response to dehydration stress. These include transformed tobacco (Brugiere *et al.*, 1999), cowpea (Figueiredo *et al.*, 2001), wheat (Nagy *et al.*, 2013) and potato when exposed to dehydration stress and high salinity conditions (Teixeira and Pereira, 2007). Interestingly, a similar result of increased protein expression by western blotting and decreased enzyme activity, as seen in tef during dehydration stress (Fig. 5.1C and 5.2D, respectively) was observed by Teixeira and Pereira (2007) in potato subjected to both drought and salt stress conditions. The authors found a decrease in GLN enzyme activity in leaves and roots of potato in response to drought and salt stress and an increase in GLN protein accumulation in potato tubers in response to drought (Teixeira and Pereira, 2007). Furthermore, the authors go on to state that the decrease in GLN enzyme activity could possibly be due to GLN enzyme inhibition or inactivation by unknown factors (Teixeira and Pereira, 2007). A similar effect has occurred in tef where certain GLN isozymes are accumulated during western blotting (Fig. 5.1C) and GLN enzyme activities are kept low or inactivated in response to dehydration stress.

5.4 Brief conclusion

In summary, the validation of a sub-set of data gathered from our statistically analysed high-abundance protein datasets (TE, TEU and MU, Chapter 3), provides an indication of the proteins and biological processes changing in response to stress conditions. These validation procedures are required to ascertain a biological change in expression data generated from iTRAQ analysis. This study attempted validation of five proteins shown to change in response to dehydration stress, through investigation of either protein accumulation using western blots or by relevant enzyme assays. The proteins included the free-radical quenching antioxidants SOD, MDHAR and POX, as well as key enzymes regulating both carbohydrate metabolism, FBA and nitrogen metabolism, GLN. Apart from GLN, the proteins investigated were shown to follow the trend predicted by iTRAQ findings, in that protein presence and activity were significantly increased in abundance, correlating with the protein datasets (TE, TEU and MU, Chapter 3).

The results from western blotting show the decrease in protein abundance of the early ROS detector, SOD and significant increase in abundance of proteins critical in the regulation of carbohydrate metabolism and nitrogen assimilation, FBA and GLN respectively, at the tested water contents of 50-55% RWC (Fig. 5.1A to C). The enzyme activities of MDHAR, FBA and POX investigated throughout dehydration stress, showed a drastic increase in activity at 60-65% RWC, with a second significant increase in activity observed for MDHAR and POX at the water contents 35-40% and 25-30% RWC, respectively (Fig. 5.2A to C). The first increase in activity at 60-65% RWC for all three enzymes, being potentially due to protection mechanisms occurring in tef at the critical water content stages above 50% RWC, just before the half-way point of dehydration (as suggested in section 2.3.2, Chapter 2). While the second increase in activity, shown towards the latter stages of dehydration between 35-40% and 25-30% RWC, for MDHAR and POX respectively, suggests at tef oxidative-stress response, where increased antioxidative proteins are stimulated in response to increasing free radicals. However, considering the loss of viability in tef plants below a RWC of 30%, accompanied by evidence of considerable membrane and subcellular damage, it is apparent that free radical damage resulting from oxidative stress (*inter alia*) is not sufficiently controlled in this species at low water contents.

Chapter 6: General discussion and conclusion

The first objective of this study was to physiologically characterise seven-week-old (pre-flowering) tef plants in response to controlled dehydration stress (in Chapter 2). These changing physiological parameters were observed and measured to establish the critical water content stages at which the damages associated with internal water loss becomes detrimental in mature tef plants such that viability is lost. The methods used to achieve these findings included monitoring changes in water content using both relative water content (RWC) and absolute water content (AWC) analysis, monitoring membrane permeability using electrolyte leakage measurements and observing the photosynthetic potential in tef using chlorophyll fluorescence analysis. Additionally, ultra-structural studies were conducted to observe the changes in sub-cellular organisation in tef leaf tissues as dehydration stress treatment proceeded and plant cell viability deteriorated.

The results show that tef has the ability to retain cellular water for up to 6 days, before internal water loss occurs. Drying over a 17 day period resulted in dehydration to approximately 30% RWC, during which viability in tef was still retained. Further imposition of dehydration stress, however, resulted in loss of viability. Damages associated with this loss were membrane rupture and consequential loss of cellular components from plant cells, as well as complete photosynthetic disruption (Figs. 2.2 and 2.3, respectively, in Chapter 2). In addition, ultra-structural studies show extensive damage to subcellular components at water contents below 30% RWC. These included compaction and distortion of cellular organelles due to a loss of cell turgor, disruption and shrinkage of vacuole structures, plasma membrane retraction and rupture, cell wall folding and eventually cell wall breakage, towards the end of dehydration treatment (~20% RWC).

A loss of water below approximately 50% RWC is critical to the survival of tef, as drastic changes in the measured physiological parameters were observed at the half-way point of dehydration (~55% RWC). At this stage, electrolyte leakage increased to a rate of 570 $\mu S.min^{-1}.gdw^{-1}$, accompanied by a progressive decline in photosynthetic potential. Although the overall trend was of the electrolyte leakage rate increasing with continuous dehydration stress, a decline in the electrolytes lost was observed in a RWC range of 40-55%. This change in membrane integrity is potentially due to induction of some protection mechanisms occurring in tef leaves in an attempt to minimise cellular water loss. Previously, Ginbot and Farrant (2011) suggested a change in cellular membrane structures in the brown-seeded tef varieties that were rearranged and repaired at rehydration upon 43% RWC, which facilitated a decline in the electrolytes lost. Thus, a certain level of protection and repair to membrane structures exists up to approximately 40% RWC in brown-seeded tef varieties. Furthermore, our results coincide with the findings previously reported by Ginbot and Farrant (2011) that some brown-seeded tef varieties (as tested here) are relatively drought tolerant, by having certain adaptive features that increase tolerance to drought

conditions to water contents above 30% RWC. Below this RWC point, however, irreversible damage occurs within tef plant cells and any previous attempts at protection against the damages associated with continuous water loss are insufficient to maintain viability. In addition, the changes in physiological measurements appear to be consistent throughout dehydration stress and further coincide with what has previously been reported for younger plants at four weeks of age (Ginbot and Farrant, 2011), indicating a level of adaptation occurring in tef plants with dehydration stress.

The second objective of this study was to conduct an in-depth proteomic analyses in tef leaf tissues during hydrated, non-stressed conditions at approximately 80% RWC and at the previously established critical water content stages in a range of 50% RWC, where tef was shown to be physiologically affected by the imposed stress conditions. To achieve this, iTRAQ mass spectrometry coupled to peptide OFFGEL fractionation and appropriate database searching with the Tef Extended and *Liliopsida* (Monocot) databases were used to observe differential regulation of tef proteins in response to dehydration stress (in Chapter 3). From the analyses, three complete dataset results were generated, the TE, TEU and MU datasets, each containing a substantial amount of database matched proteins when using the software tool PEAKS Studio 6.0.

Amongst the valid peptide-matched proteins that met FDR thresholds, a total of 5727 proteins were identified for the TE dataset, 2656 proteins identified for the TEU dataset and 4328 proteins identified for the MU dataset. Following data refinement and statistical analysis on peptide relative quantification values, it was shown that 211 proteins for the TE dataset, 111 proteins for the TEU dataset and 174 proteins for the MU dataset were differentially regulated in response to dehydration stress. Additionally, a reciprocal BLAST search through the use of OrthoMCL (Fischer *et al.*, 2011) with all three differentially regulated datasets (in the TE, TEU and MU datasets) were performed to observe common proteins and protein groups as well as to show the overlap between the three. The tool was used to establish which differentially regulated dataset (foreground) would be the most representative of the proteins changing in response to dehydration stress for further bioinformatics analyses.

From the data, the TE dataset was shown to provide the most comprehensive total protein coverage in comparison to both TEU and MU datasets when searching against the Tef Extended and *Liliopsida* databases, respectively. This could be due in part to the TE dataset being comprised of proteins detected through both uniquely matched and non-uniquely matched (shared) peptides, while the TEU and MU datasets were comprised of proteins detected only through the use of uniquely matched peptides for protein identification. For a large, non-specific and well-annotated database, such as the all monocotyledonous plants (*Liliopsida*) database, a significant amount of proteins were detected during database searching and the use of only uniquely-matched peptides for protein identification and subsequent statistical analysis was suitable (the MU dataset). However, for a much smaller, very specific

and newly sequenced genome that has been moderately-annotated such as the Tef Extended database, the use of only uniquely-matched peptides for protein identification (the TEU dataset), resulted in the number of proteins identified, being drastically reduced in comparison to the TE dataset. With the further implementation of strict data filtering to reduce noise and statistical analyses within both datasets, 111 proteins in comparison to the 211 proteins were found to be differentially regulated in the TEU and TE datasets, respectively. Thus, the TEU differentially regulated dataset was significantly lowered by 100 valid proteins meeting strict FDR thresholds through the use of only uniquely-matched peptides for protein identification.

Furthermore, a considerable amount of proteins (57 in total) within the TE differentially regulated dataset were identified as proteins generated through alternative splicing of the tef genome. This regulatory plant mechanism has been proven to enhance variation within the transcriptome and increase the functional complexity of the proteome, particularly in response to biotic and abiotic stress conditions (Duque, 2011; Staiger, 2015). It could be that many of the proteins identified as spliced variants from one pre-cursor mRNA strand, represent different protein isoforms or proteins subjected to post-translational modification (Vincent *et al.*, 2007). This would explain spliced variants such as isoforms with the same quantification values (e.g. monodehydroascorbate reductase, Table 3.1, in Chapter 3) and those few spliced variants that have differing quantification values (e.g. probable sucrose-phosphate synthase 2, Table 3.2, in Chapter 3), wherein protein degradation or post-translational modification (ubiquitination, phosphorylation etc.) could be the reason for altered quantification values.

Although software tools such as PEAKS (Ma *et al.*, 2003), used in this study and many other similar programs, allow for the input of post-translational modifications (PTMs) to search criteria (Cappadona *et al.*, 2012), it is increasingly difficult to accurately detect, differentiate and confirm these changes occurring within proteins (Parker *et al.*, 2010; Cappadona *et al.*, 2012). This is mostly due to the lack of suitable analysis tools (Cañas *et al.*, 2006). Nevertheless, proteins arising from alternative splicing should not be overlooked as these potential isoforms and altered proteins were previously shown to potentially assist in tolerance to various abiotic stresses (Abreu *et al.*, 2013), particularly in response to drought (Eckardt, 2013). Some of the studies performed in crop plants in support of these findings include: investigations in rice during drought stress and recovery (Salekdeh *et al.*, 2002), in sugar beet and maize during drought stress (Hajheidari *et al.*, 2005; Benešová *et al.*, 2012), in barley roots under saline stress (Kim *et al.*, 2005) and in wine grape cultivars subjected to both drought and increased salinity stresses (Vincent *et al.*, 2007). In tef, alternative splicing of the genome can be proposed as a regulatory mechanism that enhances adaptation to stress, by providing multiple transcripts and proteins that aid in tolerance to drought. Lastly, the TE dataset was shown to be well-represented, containing both unique and non-unique peptides with usable protein descriptions and annotations for further bioinformatics analyses, despite being generated from a newly sequenced genome (Cannarozzi *et al.*, 2014) and

moderately-annotated database. Although this is the first in-depth, exploratory study of the tef proteome and the total changes therein with dehydration stress, it is important to note that these findings would only be improved upon as more sequences are curated and annotated, opening further platforms for investigation.

To understand the biological relevance of the differentially regulated proteins in response to dehydration, functional classification, GO-term evaluation and enrichment analysis was performed using a range of bioinformatics tools. This included the use of the programs Mercator (Lohse *et al.*, 2014), MapMan (Thimm *et al.*, 2004; Usadel *et al.*, 2005), Blast2GO (Conesa *et al.*, 2005; Gotz *et al.*, 2008) and KEGG (Kanehisa and Goto, 2000) (in Chapter 4). The tools Mercator and MapMan, used in combination, were able to annotate and profile approximately half of the TE differentially regulated proteins onto known biological pathways or processes. Although these programs were able to provide some useful information for bioinformatics inference, its full potential as an annotation tool was not achieved due to many proteins being unassigned according to MapMan BIN allocations and resulted in considerable underrepresentation within annotated pathway 'maps'. Blast2GO, however, was able to classify, functionally annotate and retrieve GO-terms for more than 60% of proteins found within the TE dataset that was further used in functional enrichment analysis and interpretation of GO-terms.

While GO-term analysis of the MU dataset did not yield any functionally enriched terms, a total of 50 widely-spread GO-terms belonging to the classification ontologies CC, MF and BP, were found to be significantly enriched for proteins changing in response to dehydration stress in the TE dataset. These included GO-terms involved in biotic and abiotic stress response, signalling, transport, cellular homeostasis and pentose metabolic processes that were enriched in tef high-abundance proteins. GO-terms linked to ROS producing processes such as photosynthesis and associated light harvesting reactions as well as cell wall catabolism, manganese transport and homeostasis, the synthesis of sugars and cell wall modification, were enriched in tef low-abundance proteins. Lastly, KEGG analysis was used to observe tef proteins and enzymes mapped to various biological pathways, of which the stress responsive pathways, glutathione metabolism and ascorbate and aldarate metabolism were analysed in depth.

From the data presented in this work, an overall subtle shift in the proteome of tef occurs with dehydration stress, where proteins functioning in stress response, antioxidant protection mechanisms and those active in maintaining crucial plant cell maintenance processes are accumulated. Interestingly, abiotic stresses such as drought conditions occur in tandem with an increase in biotic stress factors, where tef showed increased susceptibility to symbiotic relationships involving parasitism and fungal responses. These results show that abiotic stress factors do not occur in isolation (Atkinson and Urwin, 2012) and that biotic stress factors should be taken into account when observing plant response to adverse changes in the environment. Furthermore, enrichment of terms associated with the decrease in abundance of

predominantly ROS-producing processes through those generated from photosynthetic reactions and metal transport were observed. This decrease in abundance levels may be in an attempt to minimise ROS proliferation associated with internal water loss.

The third objective of this study was to biologically validate proteins shown to change in quantitative expression during iTRAQ analysis and subsequent bioinformatics investigations using relevant enzyme assays and western blotting (in Chapter 5). The proteins chosen were from each of the tef high-abundance lists of proteins belonging to the TE, TEU and MU datasets. These proteins were shown to function in plant maintenance procedures during stress conditions, such as carbohydrate metabolism (FBA), nitrogen metabolism (GLN) as well as in stress protective antioxidant mechanisms against accumulative ROS (POX, MDHAR and SOD). The validation of the proteins FBA, GLN and SOD by western blotting showed a 2-fold increase in protein accumulation for FBA and GLN, while SOD displayed a significant decrease in protein accumulation in response to dehydration stress.

Validation of proteins through enzymatic assays showed the increase in activities of the antioxidant enzymes MDHAR, POX and FBA at 60-65% RWC (Fig. 5.2A to C, in Chapter 5). Increased activity of antioxidant enzymes is likely to be indicative of the enhanced oxidative stress that accompanies drought conditions (Sharma and Dubey, 2005). The increased activity of FBA, being a reversible enzyme (Uematsu *et al.*, 2012), could mean either an increase, or decrease in carbohydrate metabolism at these water contents. While carbohydrate levels were not assessed in this thesis, this change in activity correlated with changes in ultra-structural appearance of starch grains. At higher water contents, chloroplasts contained large, darkly staining starch granules, whereas at RWCs below 60%, there were numerous smaller starch granules with an electron opaque appearance (Figs. 2.4D and 2.5D, respectively, in Chapter 2). Whatever the implications thereof, this observation is indicative of sufficiently active photosynthetic processes to enable starch production; this being supported by the photosynthetic measurements presented in Chapter 2. Whether the enhanced production of carbohydrates acts as a nutrient source or aids in structural stabilisation during times of limited water availability and hence limited energy supply (Krasensky and Jonak, 2012), or perhaps functions as part of a drought tolerant adaptation within tef, would need to be investigated. This is particularly interesting since an increase in starch granules were observed in the desiccation-tolerant resurrection grass species, *E. nindensis*, when rehydrated after drying to low water contents of 10% RWC (Ginbot and Farrant, 2011).

Of the enzymes tested, GLN was shown to decrease in activity upon dehydration stress and maintain low levels throughout dehydration treatment. Although GLN protein was accumulated as indicated by western blotting at 50-55% RWC, its activity below these water contents was low (Figs. 5.1C and 5.2D, respectively, in Chapter 5), possibly being inactivated by as yet unknown factors. Some of these could potentially include regulatory PTMs acting on accumulated proteins. This result highlights the

importance of conducting enzyme assays to confirm protein presence and activity, as protein accumulation does not necessarily imply enzyme activity. Furthermore, the enzyme MDHAR, one of the key enzymes in the ascorbate-glutathione cycle and POX, an enzyme known to offer protection against the free radical H_2O_2 in plant cells, displayed a secondary increase in activity towards the latter stages of dehydration at 35-40% and 25-30% RWC, respectively (Figs. 5.2A, C, respectively). These secondary increases in activity are potentially due to stress responsive signalling, where the enzymes act in concert with enzymes of the ascorbate-glutathione cycle to neutralise excessive ROS, shown by the stress-related MapMan pathway (Fig. 4.7, in Chapter 4). However, further research into the signalling and response of enzymes in the ascorbate-glutathione cycle would need to be conducted, particularly since SOD, a key component of the ascorbate-glutathione cycle was decreased in abundance at 50-55% RWC (Fig. 5.1A) and activity of this enzyme was not investigated. In summary, the validation results were shown to support the generated iTRAQ findings and are able to provide some useful insights into the changing proteome and biochemical features of a few major biological processes occurring in tef in response to dehydration stress.

A potential limitation or weakness of the study was the use of a database generated from a newly sequenced, non-model plant genome. This resource, although representative of the plant under study in providing more protein ontological information as compared to using cross-species databases, resulted in less than expected protein identification and annotated information when investigating plant exposure to stress conditions.

6.1 Suggestions for future work

In this study, an in-depth comparative proteomics analyses of tef during hydrated (non-stressed) and dehydrated (stressed) conditions was conducted, establishing an iTRAQ pipeline for further proteomic investigation. This pipeline could be used to investigate the effect of other stresses pertinent to agricultural production of this species. For example, the effect of increased salinity can be tested in tef by observing and comparing the proteomic profiles of tef varieties shown to be salt-sensitive to those that exhibit salt-tolerance as previously reported by Asfaw and Dano (2011). However, to improve the results gained from this study and to facilitate better protein identification, a better annotated search database should be generated. To achieve this, a concatenated database consisting of both the existing tef transcriptome database and the available *Liliopsida* database, could be generated for database searching. This must be complemented with appropriate search tools to retrieve as many uniquely-scanned peptide to spectrum matches for improved protein identification (Grossmann, personal comm.). By increasing the amount of identified proteins, a larger amount of proteins could be used for ontological analyses and enrichment resulting in better inference of biological responses to dehydration. In addition, sub-cellular and sub-proteomic approaches could be explored in future where the focus could shift from total

proteome discovery to individual cellular organelles or sub-proteomes in response to stress conditions. By focussing on individual organelles and sub-proteomes, there will be better protein coverage and identification of low abundant proteins that change in response to dehydration.

This study, to our knowledge, is the first reported comparative proteomic analyses of the tef proteome in response to dehydration stress as a consequence of drought conditions and could serve as a basis for future studies and for further characterisation of tef 'omic' resources.

Reference list

Abate, E., Hussein, S., Laing, M. and Mengistu, F. 2013. Quantitative responses of tef [*Eragrostis tef* (Zucc.) Trotter] and weeping love grass [*Eragrostis curvula* (Schrad.) Nees] varieties to acid soil. *Australian Journal of Crop Science*. 7(12):1854-1860.

Abdalla, K.O. and Rafudeen, M.S. 2012. Analysis of the nuclear proteome of the resurrection plant *Xerophyta viscosa* in response to dehydration stress using iTRAQ with 2DLC and tandem mass spectrometry. *Journal of Proteomics*. 75(8):2361-2374.

Abedi, T. and Pakniyat, H. 2010. Antioxidant enzyme changes in response to drought stress in ten cultivars of oilseed Rape (*Brassica napus* L.). *Czech Journal of Genetics and Plant Breeding*. 46(1):27-34.

Abreu, I.A., Farinha, A.P., Negrão, S., Gonçalves, N., Fonseca, C., Rodrigues, M., Batista, R., Saibo, N.J.M. and Oliveira, M.M. 2013. Coping with abiotic stress: Proteome changes for crop improvement. *Journal of Proteomics*. 93:145-168.

Agarwal, P.K., Agarwal, P., Jain, P., Jha, B., Reddy, M.K. and Sopory, S.K. 2008. Constitutive overexpression of a stress-inducible small GTP-binding protein PgRab7 from *Pennisetum glaucum* enhances abiotic stress tolerance in transgenic tobacco. *Plant Cell Reports*. 27(1):105-115.

Agrawal, G.K. and Rakwal, R. 2006. Rice proteomics: A cornerstone for cereal food crop proteomes. *Mass Spectrometry Reviews*. 25(1):1-53.

Agrawal, G.K., Sarkar, A., Righetti, P.G., Pedreschi, R., Carpentier, S., Wang, T., Barkla, B.J., Kohli, A., Ndimba, B.K., Bykova, N.V., Rampitsch, C., Zolla, L., Rafudeen, M.S., Cramer, R., Bindschedler, L.V., Tsakirpaloglou, N., Ndimba, R.J., Farrant, J.M., Renaut, J., Job, D., Kikuchi, S. and Rakwal, R. 2013. A decade of plant proteomics and mass spectrometry: translation of technical advancements to food security and safety issues. *Mass Spectrometry Reviews*. 32(5):335-365.

Agrawal, G.K., Yonekura, M., Iwahashi, Y., Iwahashi, H. and Rakwal, R. 2005. System, trends and perspectives of proteomics in dicot plants: Part I: Technologies in proteome establishment. *Journal of Chromatography B*. 815(1–2):109-123.

Ahmadizadeh, M., Valizadeh, M., Zaefizadeh, M. and Shahbazi, H. 2011. Antioxidative protection and electrolyte leakage in Durum wheat under drought stress condition. *Journal of Applied Sciences Research*. 73(3):236-246.

Ahsan, N., Renaut, J. and Komatsu, S. 2009. Recent developments in the application of proteomics to the analysis of plant responses to heavy metals. *Proteomics*. 9(10):2602-2621.

Alam, I., Lee, D.G., Kim, K.H., Park, C.H., Sharmin, S.A., Lee, H., Oh, K.W., Yun, B.W. and Lee, B.H. 2010. Proteome analysis of soybean roots under waterlogging stress at an early vegetative stage. *Journal of Biosciences*. 35(1):1-15.

Alvarez, S., Berla, B.M., Sheffield, J., Cahoon, R.E., Jez, J.M. and Hicks, L.M. 2009. Comprehensive analysis of the *Brassica juncea* root proteome in response to cadmium exposure by complementary proteomic approaches. *Proteomics*. 9(9):2419-2431.

Asada, K. 2006. Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiology*. 141(2):391-396.

Asfaw, K.G. and Dano, F.I. 2011. Effects of salinity on yield and yield components of tef [*Eragrostis tef* (Zucc.) Trotter] accessions and varieties. *Current Research Journal of Biological Sciences*. 3(34):289-299.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M. and Sherlock, G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*. 25(1):25-29.

Ashraf, M. and Harris, P.J.C. 2013. Photosynthesis under stressful environments: An overview. *Photosynthetica*. 51(2):163-190.

Assefa, K., Cannarozzi, G., Girma, D., Kamies, R., Chanyalew, S., Plaza-Wüthrich, S., Blösch, R., Rindisbacher, A., Rafudeen, S. and Tadele, Z. 2015. Genetic diversity in tef [*Eragrostis tef* (Zucc.) Trotter]. *Frontiers in Plant Science*. 6(177):1-13.

Assefa, K., Ketema, S., Tefera, H., Nguyen, H., Blum, A., Ayele, M., Bai, G., Simane, B. and Kefyalew, T. 1999. Diversity among germplasm lines of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.) Trotter]. *Euphytica*. 106(1):87-97.

Assefa, K., Merker, A. and Tefera, H. 2003. Inter simple sequence repeat (ISSR) analysis of genetic diversity in tef [*Eragrostis tef* (Zucc.) Trotter]. *Hereditas*. 139(3):174-183.

Assefa, K., Tefera, H., Merker, A., Kefyalew, T. and Hundera, F. 2001. Quantitative trait diversity in tef [*Eragrostis tef* (Zucc.) Trotter] germplasm from central and northern Ethiopia. *Genetic Resources and Crop Evolution*. 48(1):53-61.

Assefa, K., Yu, J.K., Zeid, M., Belay, G., Tefera, H. and Sorrells, M.E. 2011. Breeding tef [*Eragrostis tef* (Zucc.) trotter]: conventional and molecular approaches. *Plant Breeding*. 130(1):1-9.

Atkinson, N.J. and Urwin, P.E. 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. *Journal of Experimental Botany*. 63(10):3523-43.
Ayele, M. 1999. Genetic diversity in tef [*Eragrostis tef* (Zucc.) Trotter] for osmotic adjustment, root traits, and amplified fragment length polymorphism. PhD thesis. Department of Plant and Soil Sciences. Texas Tech University, USA.
Ayele, M., Blum, A. and Nguyen, H. 2001. Diversity for osmotic adjustment and root depth in tef [*Eragrostis tef* (Zucc.) Trotter]. *Euphytica*. 121(3):237-249.
Ayele, M., Doleže, J., Van Duren, M., Brunner, H. and Zapata-Arias, F.J. 1996. Flow cytometric analysis of nuclear genome of the Ethiopian cereal tef [*Eragrostis tef* (Zucc.) Trotter]. *Genetica*. 98(2):211-215.
Baginsky, S. 2009. Plant proteomics: Concepts, applications, and novel strategies for data interpretation. *Mass Spectrometry Reviews*. 28(1):93-120.
Bai, G., Ayele, M., Tefera, H. and Nguyen, H. 2000. Genetic diversity in tef [*Eragrostis tef* (Zucc.) Trotter] and its relatives as revealed by Random Amplified Polymorphic DNAs. *Euphytica*. 112(1):15-22.
Bajji, M., Kinet, J. and Lutts, S. 2002. The use of the electrolyte leakage method for assessing cell membrane stability as a water stress tolerance test in Durum wheat. *Plant Growth Regulation*. 36(1):61-70.
Balbuena, T.S., Dias, L.L.C., Martins, M.L.B., Chiquieri, T.B., Santa-Catarina, C., Floh, E.I.S. and Silveira, V. 2011. Challenges in proteome analyses of tropical plants. *Brazilian Journal of Plant Physiology*. 23:91-104.
Balbuena, T.S., He, R., Salvato, F., Gang, D.R. and Thelen, J.J. 2012. Large-scale proteome comparative analysis of developing rhizomes of the ancient vascular plant *Equisetum hyemale*. *Frontiers in Plant Science*. 3:131.
Balsamo, R.A., Willigen, C.V., Bauer, A.M. and Farrant, J. 2006. Drought tolerance of selected *Eragrostis* species correlates with leaf tensile properties. *Annals of Botany*. 97(6):985-991.
Barkla, B.J., Vera-Estrella, R. and Pantoja, O. 2013. Progress and challenges for abiotic stress proteomics of crop plants. *Proteomics*. 13(12-13):1801-1815.
Baye, K. 2014. Teff: nutrient composition and health benefits. *In* Ethiopia Strategy Support Program. Vol. 67. Center for Food Science and Nutrition, College of Natural Sciences, Addis Ababa University, Addis Ababa, Ethiopia.
Bekele, E. 1985. A review of research on diseases of barley, tef, and wheat in Ethiopia. *In* A review of crop protection research in Ethiopia. T. Abate, editor. Proceeding of the First Ethiopian Crop Protection Symposium, Department of Crop Protection, Instituition of Agricultural Research, Addis Ababa, Ethiopia. 79-108.
Bekele, E., Fido, R.J., Tatham, A.S. and Shewry, P.R. 1995. Heterogeneity and polymorphism of seed proteins in tef (*Eragrotis tef*). *Hereditas*. 122:67-72.
Belay, G., Zemede, A., Assefa, K., Metaferia, G. and Tefera, H. 2009. Seed size effect on grain weight and agronomic performance of tef [*Eragrostis tef* (Zucc.) Trotter]. *African Journal of Agricultural Research*. 4(9):836-839.
Benešová, M., Holá, D., Fischer, L., Jedelský, P.L., Hnilička, F., Wilhelmová, N., Rothová, O., Kočová, M., Procházková, D., Honnerová, J., Fridrichová, L. and Hniličková, H. 2012. The physiology and proteomics of drought tolerance in maize: early stomatal closure as a cause of lower tolerance to short-term dehydration? *PLoS ONE*. 7(6):e38017.
Benkeblia, N. 2011. Sustainable agriculture and new biotechnologies. CRC Press, Florida, USA.
Bewley, J.D. 1979. Physiological aspects of desiccation tolerance. *Annual Review of Plant Physiology*. 30(1):195-238.
Blum, A. 1996. Crop responses to drought and the interpretation of adaptation. *Plant Growth Regulation*. 20(2):135-148.
Blum, A. and Ebercon, A. 1981. Cell membrane stability as a measure of drought and heat tolerance in wheat. *Crop Science*. 21(1):43-47.
Bluthgen, N., Brand, K., Cajavec, B., Swat, M., Herzel, H. and Beule, D. 2005. Biological profiling of gene groups utilizing Gene Ontology. *Genome Informatics*. 16(1):106-115.
Bowler, C., Montagu, M.V. and Inze, D. 1992. Superoxide dismutase and stress tolerance. *Annual Review of Plant Physiology and Plant Molecular Biology*. 43(1):83-116.
Brugiere, N., Dubois, F., Limami, A.M., Lelandais, M., Roux, Y., Sangwan, R.S. and Hirel, B. 1999. Glutamine synthetase in the phloem plays a major role in controlling proline production. *The Plant Cell*. 11(10):1995-2012.
Budak, H., Kantar, M. and Yucebilgili Kurtoglu, K. 2013. Drought tolerance in modern and wild wheat. *The Scientific World Journal*. 2013:16.
Cañas, B., López-Ferrer, D., Ramos-Fernández, A., Camafeita, E. and Calvo, E. 2006. Mass spectrometry technologies for proteomics. *Briefings in Functional Genomics & Proteomics*. 4(4):295-320.
Cannarozzi, G., Plaza-Wuthrich, S., Esfeld, K., Larti, S., Wilson, Y.S., Girma, D., de Castro, E., Chanyalew, S., Blosch, R., Farinelli, L., Lyons, E., Schneider, M., Falquet, L., Kuhlemeier, C., Assefa, K. and Tadele, Z.

2014. Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (*Eragrostis tef*). *BMC Genomics*. 15:581.
Cappadona, S., Baker, P.R., Cutillas, P.R., Heck, A.J.R. and van Breukelen, B. 2012. Current challenges in software solutions for mass spectrometry-based quantitative proteomics. *Amino Acids*. 43(3):1087-1108.
Carpentier, S.C., Coemans, B., Podevin, N., Laukens, K., Witters, E., Matsumura, H., Terauchi, R., Swennen, R. and Panis, B. 2008a. Functional genomics in a non-model crop: transcriptomics or proteomics? *Physiologia Plantarum*. 133(2):117-130.
Carpentier, S.C., Panis, B., Vertommen, A., Swennen, R., Sergeant, K., Renaut, J., Laukens, K., Witters, E., Samyn, B. and Devreese, B. 2008b. Proteome analysis of non-model plants: A challenging but powerful approach. *Mass Spectrometry Reviews*. 27(4):354-377.
Carpentier, S.C., Witters, E., Laukens, K., Van Onckelen, H., Swennen, R. and Panis, B. 2007. Banana (*Musa* spp.) as a model to study the meristem proteome: acclimation to osmotic stress. *Proteomics*. 7(1):92-105.
Chakraborty, U. and Pradhan, B. 2012. Oxidative stress in five wheat varieties (*Triticum aestivum* L.) exposed to water stress and study of their antioxidant enzyme defense system, water stress responsive metabolites and H_2O_2 accumulation. *Brazilian Journal of Plant Physiology*. 24:117-130.
Champagne, A. and Boutry, M. 2013. Proteomics of nonmodel plant species. *Proteomics*. 13(3-4):663-673.
Chance, B. and Maehly, A.C. 1955. Assay of catalase and peroxidase. *Methods in Enzymology*. 2:764-775.
Chandramouli, K. and Qian, P. Y. 2009. Proteomics: challenges, techniques and possibilities to overcome biological sample complexity. *Human Genomics and Proteomics.* 2009:239204.
Chanyalew, S., Singh, H., Tefera, H. and Sorrells, M.E. 2005. Molecular genetic map and QTL analysis of agronomic traits based on a *Eragrostis tef* x *Eragrostis pilosa* recombinant inbred population *Journal of Genetics and Breeding*. 59:53-66.
Cheung, A.Y. and De Vries, S.C. 2008. Membrane trafficking: intracellular highways and country roads. *Plant Physiology*. 147(4):1451-1453.
Chitteti, B.R. and Peng, Z. 2007. Proteome and phosphoproteome differential expression under salinity stress in rice (*Oryza sativa*) roots. *Journal of Proteome Research*. 6(5):1718-1727.
Choe, L., D'Ascenzo, M., Relkin, N.R., Pappin, D., Ross, P., Williamson, B., Guertin, S., Pribil, P. and Lee, K.H. 2007. 8-Plex Quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics*. 7(20):3651-3660.
Chrispeels, M.J., Crawford, N.M. and Schroeder, J.I. 1999. Proteins for transport of water and mineral nutrients across the membranes of plant cells. *The Plant Cell Online*. 11(4):661-675.
Chugh, V., Kaur, N. and Gupta, A.K. 2011. Evaluation of oxidative stress tolerance in maize (*Zea mays* L.) seedlings in response to drought. *Indian Journal of Biochemistry and Biophysics*. 48:47-53.
Chung, L.M., Colangelo, C.M. and Zhao, H. 2014. Data pre-processing for label-free Multiple Reaction Monitoring (MRM) experiments. *Biology (Basel)*. 3(2):383-402.
Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M. and Robles, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*. 21(18):3674-3676.
Costanza, S.H., Dewet, J.M.J. and Harlan, J.R. 1979. Literature review and numerical taxonomy of *Eragrostis tef* (tef). *Economic Botany*. 33(4):413-424.
Couee, I., Sulmon, C., Gouesbet, G. and El Amrani, A. 2006. Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. *Journal of Experimental Botany*. 57(3):449-459.
Cruz de Carvalho, M.H. 2008. Drought stress and reactive oxygen species: Production, scavenging and signaling. *Plant Signaling & Behavior*. 3(3):156-165.
CSA. 2013. Agricultural sample survey for 2012/13. Ethiopia: Statistical Bulletin Addis Ababa.
Dace, H.J., Sherwin, H.W., Illing, N. and Farrant, J.M. 1998. Use of metabolic inhibitors to elucidate mechanisms of recovery from desiccation stress in the resurrection plant *Xerophyta humilis*. *Plant Growth Regulation*. 24:171-177.
Degu, H.D. and Fujimura, T. 2010. Mapping QTLs related to plant height and root development of *Eragrostis tef* under drought. *Journal of Agricultural Science*. 2(2):62-72.
Degu, H.D., Ohta, M. and Fujimura, T. 2008. Drought tolerance of *Eragrostis tef* and development of roots. *International Journal of Plant Sciences*. 169(6):768-775.
Dionisio-Sese, M.L. and Tobita, S. 1998. Antioxidant responses of rice seedlings to salinity stress. *Plant Science*. 135(1):1-9.
Dixon, D.P., Cole, D.J. and Edwards, R. 1999. Identification and cloning of AtGST 10 (Accession Nos. AJ131580 and AJ132398), members of a novel type of plant glutathione transferases. *Plant Physiology*. 1198(1):1568-1568.
Dong, M., Gu, J., Zhang, L., Chen, P., Liu, T., Deng, J., Lu, H., Han, L. and Zhao, B. 2014. Data in support of comparative proteomics analysis of superior and inferior spikelets in hybrid rice during grain filling and response of inferior spikelets to drought stress using isobaric tags for relative and absolute quantification. *Data in Brief*. 1:51-55.

Du, Y.C., Nose, A. and Wasano, K. 1999. Effects of chilling temperature on photosynthetic rates, photosynthetic enzyme activities and metabolite levels in leaves of three sugarcane species. *Plant, Cell & Environment*. 22(3):317-324.
Duque, P. 2011. A role for SR proteins in plant stress responses. *Plant Signaling & Behavior*. 6(1):49-54.
Ebba, T. 1975. Tef (*Eragrostis tef*) cultivars: Morphology and Classification. Part II. Expt. Sta. Bull. 66. Addis Ababa University, College of Agriculture, Dire Dawa, Ethiopia.
Eckardt, N.A. 2013. Alternative splicing confers a dual role in polar auxin transport and drought stress tolerance to the major facilitator superfamily transporter ZIFL1. *The Plant Cell*. 25(3):779.
El-Alfy, T.S., Ezzat, S.M. and Sleem, A.A. 2012. Chemical and biological study of the seeds of *Eragrostis tef* (Zucc.) Trotter. *Natural Product Research*. 26(7):619-629.
Enright, A.J., Van Dongen, S. and Ouzounis, C.A. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research*. 30(7):1575-1584.
Eugeni Piller, L., Besagni, C., Ksas, B., Rumeau, D., Brehelin, C., Glauser, G., Kessler, F. and Havaux, M. 2011. Chloroplast lipid droplet type II NAD(P)H quinone oxidoreductase is essential for prenylquinone metabolism and vitamin K1 accumulation. *Proceedings of the National Academy of Sciences of the United States of America* 108(34):14354-14359.
Farrant, J.M. 2000. A comparison of mechanisms of desiccation tolerance among three angiosperm resurrection plant species. *Plant Ecology*. 151(1):29-39.
Farrant, J.M., Brandt, W. and Lindsey, G.G. 2007. An overview of mechanisms of desiccation tolerance in selected angiosperm resurrection plants. *Plant Stress*. 1(1):72-84.
Farrant, J.M., Vander Willigen, C., Loffell, D.A., Bartsch, S. and Whittaker, A. 2003. An investigation into the role of light during desiccation of three angiosperm resurrection plants. *Plant, Cell and Environment*. 26:1275–1286.
Fercha, A., Capriotti, A.L., Caruso, G., Cavaliere, C., Samperi, R., Stampachiacchiere, S. and Lagana, A. 2014. Comparative analysis of metabolic proteome variation in ascorbate-primed and unprimed wheat seeds during germination under salt stress. *Journal of Proteomics*. 108:238-257.
Figueiredo, M.d.V.B., Bezerra-Neto, E. and Burity, H.A. 2001. Water stress response on the enzymatic activity in cowpea nodules. *Brazilian Journal of Microbiology*. 32:195-200.
Fischer, S., Brunk, B.P., Chen, F., Gao, X., Harb, O.S., Iodice, J.B., Shanmugam, D., Roos, D.S. and Stoeckert, C.J., Jr. 2011. Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups. *Current Protocols in Bioinformatics*. Chapter 6:Unit 6.12.11-19.
Fleury, D., Jefferies, S., Kuchel, H. and Langridge, P. 2010. Genetic and genomic tools to improve drought tolerance in wheat. *Journal of Experimental Botany*. 61(12):3211-3222.
Ford, K.L., Cassin, A. and Bacic, A. 2011. Quantitative proteomic analysis of wheat cultivars with differing drought stress tolerance. *Frontiers in Plant Science*. 2:44.
Foyer, C.H., Lelandais, M. and Kunert, K.J. 1994. Photooxidative stress in plants. *Physiologia Plantarum*. 92(4):696-717.
Foyer, C.H. and Noctor, G. 2009. Redox regulation in photosynthetic organisms: signaling, acclimation, and practical implications. *Antioxidants and Redox Signalling*. 11(4):861-905.
Fukai, S. and Cooper, M. 1995. Development of drought-resistant cultivars using physiomorphological traits in rice. *Field Crops Research*. 40(2):67-86.
Gan, C.S., Chong, P.K., Pham, T.K. and Wright, P.C. 2007. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *Journal of Proteome Research*. 6(2):821-827.
Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y. and Zhang, J. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*. 5(10):R80.
Ghosh, D. and Xu, J. 2014. Abiotic stress responses in plant roots: a proteomics perspective. *Frontiers in Plant Science*. 5(6):1-13.
Gill, S.S. and Tuteja, N. 2010. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiology and Biochemistry*. 48(12):909-930.
Ginbot, Z. and Farrant, J.M. 2011. Physiological response of selected *Eragrostis* species to water-deficit stress. *African Journal of Biotechnology*. 10(51):10405-10417.
Girma, D., Assefa, K., Chanyalew, S., Cannarozzi, G., Kuhlemeier, C. and Tadele, Z. 2014. The origins and progress of genomics research on tef (*Eragrostis tef*). *Plant Biotechnology Journal*. 12(5):534-540.
Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J. and Conesa, A. 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*. 36(10):3420-3435.

Graciet, E., Lebreton, S. and Gontero, B. 2004. Emergence of new regulatory mechanisms in the Benson–Calvin pathway via protein–protein interactions: a glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase complex. *Journal of Experimental Botany*. 55(400):1245-1254.
Grene, R. 2002. Oxidative stress and acclimation mechanisms in plants. *The Arabidopsis Book / American Society of Plant Biologists*. 1:e0036.
Grossmann, J., Fischer, B., Baerenfaller, K., Owiti, J., Buhmann, J.M., Gruissem, W. and Baginsky, S. 2007. A workflow to increase the detection rate of proteins from unsequenced organisms in high-throughput proteomics experiments. *Proteomics*. 7(23):4245-4254.
Gupta, N. and Pevzner, P.A. 2009. False discovery rates of protein identifications: a strike against the two-peptide rule. *Journal of Proteome Research*. 8(9):4173-4181.
Ha, S.B., Smith, A.P., Howden, R., Dietrich, W.M., Bugg, S., O'Connell, M.J., Goldsbrough, P.B. and Cobbett, C.S. 1999. Phytochelatin synthase genes from *Arabidopsis* and the yeast *Schizosaccharomyces pombe*. *The Plant Cell*. 11(6):1153-1164.
Hajheidari, M., Abdollahian-Noghabi, M., Askari, H., Heidari, M., Sadeghian, S.Y., Ober, E.S. and Salekdeh, G.H. 2005. Proteome analysis of sugar beet leaves under drought stress. *Proteomics*. 5(4):950-960.
Hashiguchi, A., Ahsan, N. and Komatsu, S. 2010. Proteomics application of crops in the context of climatic changes. *Food Research International*. 43(7):1803-1813.
Hiremath, P.J., Farmer, A., Cannon, S.B., Woodward, J., Kudapa, H., Tuteja, R., Kumar, A., BhanuPrakash, A., Mulaosmanovic, B., Gujaria, N., Krishnamurthy, L., Gaur, P.M., KaviKishor, P.B., Shah, T., Srinivasan, R., Lohse, M., Xiao, Y., Town, C.D., Cook, D.R., May, G.D. and Varshney, R.K. 2011. Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. *Plant Biotechnology Journal*. 9(8):922-931.
Hoogenboom, G., Huck, M.G. and Peterson, C.M. 1987. Root growth rate of soybean as affected by drought stress. *Agronomy Journal*. 79(4):607-614.
Hu, H., Dai, M., Yao, J., Xiao, B., Li, X., Zhang, Q. and Xiong, L. 2006. Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proceedings of the National Academy of Sciences of the United States of America*. 103(35):12987-12992.
Huang, G.J., Deng, J.S., Chen, H.J., Huang, S.S., Shih, C.C. and Lin, Y.H. 2013. Dehydroascorbate reductase and monodehydroascorbate reductase activities of two metallothionein-like proteins from sweet potato (*Ipomoea batatas* [L.] Lam. 'Tainong 57') storage roots. *Botanical Studies*. 54(1):7.
Huda, K.M.K., Banu, M.S.A., Tuteja, R. and Tuteja, N. 2013. Global calcium transducer P-type Ca^{2+}-ATPases open new avenues for agriculture by regulating stress signalling. *Journal of Experimental Botany*. 64(11):3099-3109.
Ingram, A.L. and Doyle, J.J. 2003. The origin and evolution of *Eragrostis tef* (Poaceae) and related polyploids: evidence from nuclear waxy and plastid rps16. *American Journal of Botany*. 90(1):116-122.
International Rice Genome Sequencing Project. 2005. The map-base sequence of the rice genome. *Nature.* 436(7052):793-800.
Isaacson, T., Damasceno, C.M., Saravanan, R.S., He, Y., Catala, C., Saladie, M. and Rose, J.K. 2006. Sample extraction techniques for enhanced proteomic analysis of plant tissues. *Nature Protocols*. 1(2):769-774.
Jacoby, R.P., Li, L., Huang, S., Pong Lee, C., Millar, A.H. and Taylor, N.L. 2012. Mitochondrial composition, function and stress response in plants. *Journal of Integrative Plant Biology*. 54(11):887-906.
Jain, M. 2013. Emerging role of metabolic pathways in abiotic stress tolerance. *Journal of Plant Biochemistry and Physiology*. 1:108.
Jedmowski, C., Ashoub, A., Beckhaus, T., Berberich, T., Karas, M. and Bruggemann, W. 2014. Comparative analysis of *Sorghum bicolor* proteome in response to drought stress and following recovery. *International Journal of Proteomics*. 2014:10.
Jiang, S.S., Liang, X.N., Li, X., Wang, S.L., Lv, D.W., Ma, C.Y., Li, X.H., Ma, W.J. and Yan, Y.M. 2012. Wheat drought-responsive grain proteome analysis by linear and nonlinear 2-DE and MALDI-TOF mass spectrometry. *International Journal of Molecular Sciences*. 13(12):16065-16083.
Jimenez, M.S., Gonzalez-Rodriguez, A.M., Morales, D., Cid, M.C., Socorro, A.R. and Caballero, M. 1997. Evaluation of chlorophyll fluorescence as a tool for salt stress detection in roses. *Photosynthetica*. 33(2):291-301.
Johnson, S., Lim, F.L., Finkler, A., Fromm, H., Slabas, A. and Knight, M. 2014. Transcriptomic analysis of *Sorghum bicolor* responding to combined heat and drought stress. *BMC Genomics*. 15(1):456.
Jones, B.M.G., Ponti, J., Tavassoli, A. and Dixon, P.A. 1978. Relationships of the Ethiopian cereal tef (*Eragrostis tef* (Zucc.) Trotter): evidence from morphology and chromosome number. *Annals of Botany*. 42(6):1369-1373.
Jorrin-Novo, J.V., Maldonado, A.M., Echevarria-Zomeno, S., Valledor, L., Castillejo, M.A., Curto, M., Valero, J., Sghaier, B., Donoso, G. and Redondo, I. 2009. Plant proteomics update (2007-2008): Second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards,

increase plant proteome coverage and expand biological knowledge. *Journal of Proteomics*. 72(3):285-314.
Jöst, M., Esfeld, K., Burian, A., Cannarozzi, G., Chanyalew, S., Kuhlemeier, C., Assefa, K. and Tadele, Z. 2014. Semi-dwarfism and lodging tolerance in tef (*Eragrostis tef*) is linked to a mutation in the α-Tubulin 1 gene. *Journal of Experimental Botany*.
Kadioglu, A., Terzi, R., Saruhan, N. and Saglam, A. 2012. Current advances in the investigation of leaf rolling caused by biotic and abiotic stress factors. *Plant Science*. 182:42-48.
Kamal, A.H.M., Cho, K., Kim, D.E., Uozumi, N., Chung, K.Y., Lee, S.Y., Choi, J.S., Cho, S.W., Shin, C.S. and Woo, S.H. 2012. Changes in physiology and protein abundance in salt-stressed wheat chloroplasts. *Molecular Biology Reports*. 39(9):9059-9074.
Kambiranda, D., Katam, R., Basha, S.M. and Siebert, S. 2013. iTRAQ-based quantitative proteomics of developing and ripening muscadine grape berry. *Journal of Proteome Research*. 13(2):555-569.
Kamies, R., Rafudeen, M.S. and Farrant, J.M. 2010. The use of aeroponics to investigate antioxidant activity in the roots of *Xerophyta viscosa*. *Plant Growth Regulation*. 62(3):203-211.
Kanehisa, M. and Goto, S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*. 28(1):27-30.
Karp, N.A., Huber, W., Sadowski, P.G., Charles, P.D., Hester, S.V. and Lilley, K.S. 2010. Addressing accuracy and precision issues in iTRAQ quantitation. *Molecular and Cellular Proteomics*. 9(9):1885-1897.
Kazan, K. 2003. Alternative splicing and proteome diversity in plants: the tip of the iceberg has just emerged. *Trends in Plant Science*. 8(10):468-471.
Kessner, D., Chambers, M., Burke, R., Agus, D. and Mallick, P. 2008. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics*. 24(21):2534-2536.
Ketema, S. 1997. Tef, *Eragrostis tef* (Zucc.) Trotter. Institute of Plant Genetics and Crop Plant Research, Gatersleben/International Plant Genetic Resources Institute, Rome, Italy 52 pp.
Keunen, E., Remans, T., Bohler, S., Vangronsveld, J. and Cuypers, A. 2011. Metal-induced oxidative stress and plant mitochondria. *International Journal of Molecular Sciences*. 12(10):6894-6918.
Kim, S.T., Kim, S.G., Agrawal, G.K., Kikuchi, S. and Rakwal, R. 2014. Rice proteomics: a model system for crop improvement and food security. *Proteomics*. 14(4-5):593-610.
Kim, S.Y., Lim, J.H., Park, M.R., Kim, Y.J., Park, T.I., Seo, Y.W., Choi, K.G. and Yun, S.J. 2005. Enhanced antioxidant enzymes are associated with reduced hydrogen peroxide in barley roots under saline stress. *Journal of Biochemistry and Molecular Biology*. 38(2):218-224.
Kingston-Smith, A.H. and Foyer, C.H. 2000. Overexpression of Mn-superoxide dismutase in maize leaves leads to increased monodehydroascorbate reductase, dehydroascorbate reductase and glutathione reductase activities. *Journal of Experimental Botany*. 51(352):1867-1877.
Kocheva, K., Lambrev, P., Georgiev, G., Goltsev, V. and Karabaliev, M. 2004. Evaluation of chlorophyll fluorescence and membrane injury in the leaves of barley cultivars under osmotic stress. *Bioelectrochemistry*. 63(1-2):121-124.
Komatsu, S., Kamal, A.H.M. and Hossain, Z. 2014. Wheat proteomics: Proteome modulation and abiotic stress acclimation. *Frontiers in Plant Science*. 5.
Komatsu, S., Kobayashi, Y., Nishizawa, K., Nanjo, Y. and Furukawa, K. 2010. Comparative proteomics analysis of differentially expressed proteins in soybean cell wall during flooding stress. *Amino Acids*. 39(5):1435-1449.
Kornblum, H.I. 2007. Introduction to neural stem cells. *Stroke*. 38(2):810-816.
Kosová, K., Vítámvás, P., Prášil, I.T. and Renaut, J. 2011. Plant proteome changes under abiotic stress-contribution of proteomics studies to understanding plant stress response. *Journal of Proteomics*. 74(8):1301-1322.
Kranner, I., Birtic, S., Anderson, K.M. and Pritchard, H.W. 2006. Glutathione half-cell reduction potential: a universal stress marker and modulator of programmed cell death? *Free Radical Biological Medicine*. 40(12):2155-2165.
Krasensky, J. and Jonak, C. 2012. Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. *Journal of Experimental Botany*.
Kreitschitz, A., Tadele, Z. and Gola, E.M. 2009. Slime cells on the surface of *Eragrostis* seeds maintain a level of moisture around the grain to enhance germination. *Seed Science Research* 19:27-35.
Kumar, R., Kumar, A., Subba, P., Gayali, S., Barua, P., Chakraborty, S. and Chakraborty, N. 2014. Nuclear phosphoproteome of developing chickpea seedlings (*Cicer arietinum* L.) and protein-kinase interaction network. *Journal of Proteomics*. 105:58-73.
Lan, P., Li, W., Wen, T.N., Shiau, J.Y., Wu, Y.C., Lin, W. and Schmidt, W. 2011. iTRAQ protein profile analysis of *Arabidopsis* roots reveals new aspects critical for iron homeostasis. *Plant Physiology*. 155(2):821-834.
Lester, R.N. and Bekele, E. 1981. Amino acid composition of the cereal tef and related species of *Eragrostis* (Gramineae). *Cereal Chemistry*. 58(2):113-115.

Levitt, J. 1980. Responses of plants to environmental stresses. Vol.2. Water, radiation, salt and other stresses. Academic Press, New York.
Lin, Z., Zhang, X., Yang, X., Li, G., Tang, S., Wang, S., Ding, Y. and Liu, Z. 2014. Proteomic analysis of proteins related to rice grain chalkiness using iTRAQ and a novel comparison system based on a notched-belly mutant with white-belly. *BMC Plant Biology*. 14(1):163.
Liu, F., Guo, J., Bai, P., Duan, Y., Wang, X., Cheng, Y., Feng, H., Huang, L. and Kang, Z. 2012. Wheat *TaRab7* GTPase is part of the signaling pathway in responses to stripe rust and abiotic stimuli. *PLoS ONE*. 7(5):e37146.
Liu, G.T., Ma, L., Duan, W., Wang, B.C., Li, J.H., Xu, H.G., Yan, X.Q., Yan, B.F., Li, S.H. and Wang, L.J. 2014. Differential proteomic analysis of grapevine leaves by iTRAQ reveals responses to heat stress and subsequent recovery. *BMC Plant Biology*. 14(1):110.
Lohse, M., Nagel, A., Herter, T., May, P., Schroda, M., Zrenner, R., Tohge, T., Fernie, A.R., Stitt, M. and Usadel, B. 2014. Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. *Plant Cell and Environment*. 37(5):1250-1258.
Ludlow, M.M. and Muchow, R.C. 1990. A critical evaluation of traits for improving crop yields in water-limited environments. *Advances in Agronomy*:107-153.
Ma, B., Zhang, K., Hendrie, C., Liang, C., Li, M., Doherty-Kirby, A. and Lajoie, G. 2003. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*. 17(20):2337-2342.
Machado, A.T., Sodek, L., Paterniani, E. and Fernandes, M.S. 2001. Nitrate reductase and glutamine synthetase activities in S1 endogamic families of the maize populations Sol da Manhã NF and Catetão. *Revista Brasileira de Fisiologia Vegetal*. 13:88-102.
Mamo, T. and Parsons, J.W. 1987. Iron nutrition of *Eragrostis tef* (tef). *Tropical Agriculture*. 64:313-317.
Manmathan, H., Shaner, D., Snelling, J., Tisserat, N. and Lapitan, N. 2013. Virus-induced gene silencing of *Arabidopsis thaliana* gene homologues in wheat identifies genes conferring improved drought tolerance. *Journal of Experimental Botany*. 64(5):1381-1392.
Marshall, J.G. and Dumbroff, E.B. 1999. Turgor regulation via cell wall adjustment in white spruce. *Plant Physiology*. 119(1):313-320.
Martinez-Esteso, M., Vilella-Anton, M., Pedreno, M., Valero, M. and Bru-Martinez, R. 2013. iTRAQ-based protein profiling provides insights into the central metabolism changes driving grape berry development and ripening. *BMC Plant Biology*. 13(1):167.
Martinez-Esteso, M.J., Casado-Vela, J., Selles-Marchart, S., Pedreno, M.A. and Bru-Martinez, R. 2014. Differential plant proteome analysis by isobaric tags for relative and absolute quantitation (iTRAQ). *Methods in Molecular Biology*. 1072:155-169.
Maxwell, K. and Johnson, G.N. 2000. Chlorophyll fluorescence - a practical guide. *Journal of Experimental Botany*. 51(345):659-668.
McCord, J.M. and Fridovich, I. 1969. Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). *Journal of Biological Chemistry*. 244(22):6049-6055.
Mengistu, D.K. 2009. The influence of soil water deficit imposed during various developmental phases on physiological processes of tef (*Eragrostis tef*). *Agriculture, Ecosystems & Environment*. 132(3–4):283-289.
Mengistu, D.K. and Mekonnen, L.S. 2012. Integrated agronomic crop managements to improve tef productivity under terminal drought. *In* Water Stress. I.M.M. Rahman, editor. INTECH Open Access Publisher. 235-255.
Merrill, S.D. and Rawlins, S.L. 1979. Distribution and growth of sorghum roots in response to irrigation frequency. *Agronomy Journal*. 71(5):738-745.
Miflin, B.J. and Habash, D.Z. 2002. The role of glutamine synthetase and glutamate dehydrogenase in nitrogen assimilation and possibilities for improvement in the nitrogen utilization of crops. *Journal of Experimental Botany*. 53(370):979-987.
Mittler, R., Vanderauwera, S., Gollery, M. and Van Breusegem, F. 2004. Reactive oxygen gene network of plants. *Trends in Plant Science*. 9(10):490-498.
Miyake, C. and Asada, K. 1992. Thylakoid-bound ascorbate peroxidase in spinach chloroplasts and photoreduction of its primary oxidation product monodehydroascorbate radicals in thylakoids. *Plant and Cell Physiology*. 33(5):541-553.
Mochida, K. and Shinozaki, K. 2010. Genomics and bioinformatics resources for crop improvement. *Plant and Cell Physiology*. 51(4):497-523.
Molaei, P., Ebadi, A., Namvar, A. and Bejandi, T.K. 2012. Water relation, solute accumulation and cell membrane injury in sesame (*Sesamum indicum* L.) cultivars subjected to water stress. *Annals of Biological Research*. 3(4):1833-1838.

Molina-Rueda, J.J., Tsai, C.J. and Kirby, E.G. 2013. The *Populus* superoxide dismutase gene family and its responses to drought stress in transgenic poplar overexpressing a pine cytosolic glutamine synthetase (GS1a). *PLoS ONE*. 8(2):e56421.
Moore, J.P., Nguema-Ona, E., Chevalier, L., Lindsey, G.G., Brandt, W.F., Lerouge, P., Farrant, J.M. and Driouich, A. 2006. Response of the leaf cell wall to desiccation in the resurrection plant *Myrothamnus flabellifolius*. *Plant Physiology*. 141(2):651-662.
Moore, J.P., Vicre-Gibouin, M., Farrant, J.M. and Driouich, A. 2008. Adaptations of higher plant cell walls to water loss: drought vs desiccation. *Physiologia Plantarum*. 134(2):237-245.
Morell, S., Follmann, H., De Tullio, M. and Häberlein, I. 1997. Dehydroascorbate and dehydroascorbate reductase are phantom indicators of oxidative stress in plants. *FEBS letters*. 414(3):567-570.
Mundree, S.G., Whittaker, A., Thomson, J.A. and Farrant, J.M. 2000. An aldose reductase homolog from the resurrection plant *Xerophyta viscosa* Baker. *Planta*. 211(5):693-700.
Murthy, S.M., Devaraj, V.R., Anitha, P. and Tejavathi, D.H. 2012. Studies on the activities of antioxidant enzymes under induced drought stress in *in vivo* and *in vitro* plants of *Macrotyloma uniflorum* (Lam.) Verdc. . *Recent Research in Science and Technology*. 4(2):34-37.
Nagy, Z., Nemeth, E., Guoth, A., Bona, L., Wodala, B. and Pecsvaradi, A. 2013. Metabolic indicators of drought stress tolerance in wheat: glutamine synthetase isoenzymes and rubisco. *Plant Physiology and Biochemistry*. 67:48-54.
Nakashima, K., Kiyosue, T., Yamaguchi-Shinozaki, K. and Shinozaki, K. 1997. A nuclear gene, *ERD1*, encoding a chloroplast-targeted Clp protease regulatory subunit homolog is not only induced by water stress but also developmentally up-regulated during senescence in *Arabidopsis thaliana*. *The Plant Journal*. 12(4):851-861.
Nanjo, Y., Nouri, M.Z. and Komatsu, S. 2011. Quantitative proteomic analyses of crop seedlings subjected to stress conditions; a commentary. *Phytochemistry*. 72(10):1263-1272.
Nanjo, Y., Skultety, L., Ashraf, Y. and Komatsu, S. 2010. Comparative proteomic analysis of early-stage soybean seedlings responses to flooding by using gel and gel-free techniques. *Journal of Proteome Research*. 9(8):3989-4002.
Navari-Izzo, F., Meneguzzo, S., Loggini, B., Vazzana, C. and Sgherri, C.L.M. 1997. The role of the glutathione system during dehydration of *Boea hygroscopica*. *Physiologia Plantarum*. 99(1):23-30.
Nazarli, H., Zardashti, M.R., Darvishzadeh, R. and Mohammadi, M. 2011. Change in activity of antioxidative enzymes in young leaves of sunflower (*Helianthus annuus* L.) by application of super absorbent synthetic polymers under drought stress condition. *Australian Journal of Crop Science*. 5(11):1334-1338.
Neilson, K.A., Gammulla, C.G., Mirzaei, M., Imin, N. and Haynes, P.A. 2010. Proteomic analysis of temperature stress in plants. *Proteomics*. 10(4):828-845.
Neilson, K.A., Mariani, M. and Haynes, P.A. 2011. Quantitative proteomic analysis of cold-responsive proteins in rice. *Proteomics*. 11(9):1696-1706.
Nestler, J., Schutz, W. and Hochholdinger, F. 2011. Conserved and unique features of the maize (*Zea mays* L.) root hair proteome. *Journal of Proteome Research*. 10(5):2525-2537.
Noctor, G., Arisi, A.C.M., Jouanin, L., Kunert, K.J., Rennenberg, H. and Foyer, C.H. 1998. Glutathione: biosynthesis, metabolism and relationship to stress tolerance explored in transformed plants. *Journal of Experimental Botany*. 49(321):623-647.
Noctor, G. and Foyer, C.H. 1998. Ascorbate and glutathione: keeping active oxygen under control. *Annual Review of Plant Physiology and Plant Molecular Biology*. 49:249-279.
O'Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *The Journal of Biological Chemistry*. 250(10):4007-4021.
Oh, M.W., Roy, S.K., Kamal, A.H., Cho, K., Cho, S.W., Park, C.S., Choi, J.S., Komatsu, S. and Woo, S.H. 2014. Proteome analysis of roots of wheat seedlings under aluminum stress. *Molecular Biology Reports*. 41(2):671-681.
Oh, S.J., Song, S.I., Kim, Y.S., Jang, H.J., Kim, S.Y., Kim, M., Kim, Y.K., Nahm, B.H. and Kim, J.K. 2005. Arabidopsis *CBF3/DREB1A* and *ABF3* in transgenic rice increased tolerance to abiotic stress without stunting growth. *Plant Physiology*. 138(1):341-351.
Ong, S.E. and Mann, M. 2005. Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology*. 1(5):252-262.
Ow, S.Y., Salim, M., Noirel, J., Evans, C., Rehman, I. and Wright, P.C. 2009. iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". *Journal of Proteome Research*. 8(11):5347-5355.
Owiti, J., Grossmann, J., Gehrig, P., Dessimoz, C., Laloi, C., Hansen, M.B., Gruissem, W. and Vanderschuren, H. 2011. iTRAQ-based analysis of changes in the cassava root proteome reveals pathways associated with post-harvest physiological deterioration. *Plant Journal*. 67(1):145-156.
Pandey, A. and Mann, M. 2000. Proteomics to study genes and genomes. *Nature*. 405(6788):837-846.

Pandey, A., Chakraborty, S., Datta, A. and Chakraborty, N. 2008. Proteomics approach to identify dehydration responsive nuclear proteins from chickpea (*Cicer arietinum* L.). *Molecular and Cellular Proteomics*. 7(1):88-107.
Panse, C. and Grossmann, J. 2012. protViz: visualizing and analyzing mass spectrometry related data in proteomics. *R-package* version 0.1.26.
Parker, C.E., Mocanu, V., Mocanu, M., Dicheva, N. and Warren, M.R. 2010. Mass spectrometry for post-translational modifications. *In* Neuroproteomics. O. Alzate, editor. CRC Press, Boca Raton (FL). Chapter 6
Patterson, J., Ford, K., Cassin, A., Natera, S. and Bacic, A. 2007. Increased abundance of proteins involved in phytosiderophore production in boron-tolerant barley. *Plant Physiology*. 144(3):1612-1631.
Patterson, S.D. and Aebersold, R.H. 2003. Proteomics: the first decade and beyond. *Nature Genetics*. 33:311-323.
Peiter, E., Montanini, B., Gobert, A., Pedas, P., Husted, S., Maathuis, F.J.M., Blaudez, D., Chalot, M. and Sanders, D. 2007. A secretory pathway-localized cation diffusion facilitator confers plant manganese tolerance. *Proceedings of the National Academy of Sciences of the United States of America*. 104(20):8532-8537.
Pelleschi, S., Rocher, J.P. and Prioul, J.L. 1997. Effect of water restriction on carbohydrate metabolism and photosynthesis in mature maize leaves. *Plant, Cell & Environment*. 20(4):493-503.
Peltier, G. and Cournac, L. 2002. Chlororespiration. *Annual Review of Plant Biology*. 53:523-550.
Peng, L., Yamamoto, H. and Shikanai, T. 2011. Structure and biogenesis of the chloroplast NAD(P)H dehydrogenase complex. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*. 1807(8):945-953.
Plaza-Wüthrich, S. and Tadele, Z. 2012. Millet improvement through regeneration and transformation *Biotechnology and Molecular Biology Review*. 7(2):48-61.
Premachandra, G. and Shimada, T. 1987. The measurement of cell membrane stability using polyethylene glycol as a drought tolerance test in wheat. *Japanese Journal of Crop Science*. 56(1):92-98.
Primmer, C.R., Papakostas, S., Leder, E.H., Davis, M.J. and Ragan, M.A. 2013. Annotated genes and nonannotated genomes: cross-species use of Gene Ontology in ecology and evolution research. *Molecular Ecology*. 22(12):3216-3241.
Qin, J., Gu, F., Liu, D., Yin, C., Zhao, S., Chen, H., Zhang, J., Yang, C., Zhan, X. and Zhang, M. 2013. Proteomic analysis of elite soybean Jidou17 and its parents using iTRAQ-based quantitative approaches. *Proteome Science*. 11(1):1-11.
Quan, L.J., Zhang, B., Shi, W.W. and Li, H.Y. 2008. Hydrogen peroxide in plants: a versatile molecule of the reactive oxygen species network. *Journal of Integrative Plant Biology*. 50(1):2-18.
Quaye, A.K., Laryea, K.B. and Abeney-Mickson, S. 2009. Soil water and nitrogen interaction effects on maize (*Zea mays* L.) grown on a vertisol *Journal of Forestry, Horticulture and Soil Science*. 3(1):1-11.
Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. and Lopez, R. 2005. InterProScan: protein domains identifier. *Nucleic Acids Research*. 33(Web Server issue):W116-W120.
Qureshi, M.I., Qadir, S. and Zolla, L. 2007. Proteomics-based dissection of stress-responsive pathways in plants. *Journal of Plant Physiology*. 164(10):1239-1260.
Reddy, A.R., Chaitanya, K.V. and Vivekanandan, M. 2004. Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. *Journal of Plant Physiology*. 161(11):1189-1202.
Reddy, A.S. 2007. Alternative splicing of pre-messenger RNAs in plants in the genomic era. *Annual Review of Plant Biology*. 58:267-294.
Rhodes, D., Rendon, G.A. and Stewart, G.R. 1975. The control of glutamine synthetase level in *Lemna minor* L. *Planta*. 125(3):201-211.
Rolny, N., Costa, L., Carrión, C. and Guiamet, J.J. 2011. Is the electrolyte leakage assay an unequivocal test of membrane deterioration during leaf senescence? *Plant Physiology and Biochemistry*. 49(10):1220-1227.
Romero-Rodriguez, M.C., Pascual, J., Valledor, L. and Jorrin-Novo, J. 2014. Improving the quality of protein identification in non-model species. Characterization of *Quercus ilex* seed and *Pinus radiata* needle proteomes by using SEQUEST and custom databases. *Journal of Proteomics*. 105:85-91.
Ross, P.L., Huang, Y.N., Marchese, J.N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S., Purkayastha, S., Juhasz, P., Martin, S., Bartlet-Jones, M., He, F., Jacobson, A. and Pappin, D.J. 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular and Cellular Proteomics*. 3(12):1154-1169.
Rotter, A., Camps, C., Lohse, M., Kappel, C., Pilati, S., Hren, M., Stitt, M., Coutos-Thevenot, P., Moser, C., Usadel, B., Delrot, S. and Gruden, K. 2009. Gene expression profiling in susceptible interaction of grapevine with its fungal pathogen *Eutypa lata*: Extending MapMan ontology for grapevine. *BMC Plant Biology*. 9(1):104.
Rumeau, D., Becuwe-Linka, N., Beyly, A., Louwagie, M., Garin, J. and Peltier, G. 2005. New subunits NDH-M, -N, and -O, encoded by nuclear genes, are essential for plastid Ndh complex functioning in higher plants. *Plant Cell*. 17(1):219-232.
Salekdeh, G.H. and Komatsu, S. 2007. Crop proteomics: Aim at sustainable agriculture of tomorrow. *Proteomics*. 7(16):2976-2996.

Salekdeh, G.H., Siopongco, J., Wade, L.J., Ghareyazie, B. and Bennett, J. 2002. Proteomic analysis of rice leaves during drought stress and recovery. *Proteomics*. 2(9):1131-1145.
Sarieva, G.E., Kenzhebaeva, S.S. and Lichtenthaler, H.K. 2010. Adaptation potential of photosynthesis in wheat cultivars with a capability of leaf rolling under high temperature conditions. *Russian Journal of Plant Physiology*. 57(1):28-36.
Saruhan, N., Terzi, R., Saglam, A. and Kadioglu, A. 2009. The relationship between leaf rolling and ascorbate-glutathione cycle enzymes in apoplastic and symplastic areas of *Ctenanthe setosa* subjected to drought stress. *Biological Research*. 42(3):315-326.
Schlede, H. 1989. Distribution of acid soils and liming materials in Ethiopia. Ethiopian Institute of Geological Survey, Ministry of Mines and Energy, Addis Ababa, Ethiopia.
Shah, K., Kumar, R.G., Verma, S. and Dubey, R.S. 2001. Effect of cadmium on lipid peroxidation, superoxide anion generation and activities of antioxidant enzymes in growing rice seedlings. *Plant Science*. 161(6):1135-1144.
Sharma, P. and Dubey, R. 2005. Drought induces oxidative stress and enhances the activities of antioxidant enzymes in growing rice seedlings. *Plant Growth Regulation*. 46(3):209-221.
Sherwin, H.W. and Farrant, J.M. 1996. Differences in rehydration of three desiccation-tolerant angiosperm species. *Annals of Botany*. 78:703-710.
Shiferaw, B. and Baker, D.A. 1996. An evaluation of drought screening techniques for *Eragrostis tef*. *Tropical Science*. 36:74-85.
Shiferaw, W., Balcha, A. and Mohammed, H. 2012a. Evaluation of drought tolerance indices in tef [*Eragrostis tef* (Zucc.)Trotter] *African Journal of Agricultural Research*. 7(23):3433-3438.
Shiferaw, W., Balcha, A. and Mohammed, H. 2012b. Genetic variation for grain yield and yield related traits in tef [*Eragrostis tef* (Zucc.)Trotter] under moisture stress and non-stress environments. *American Journal of Plant Sciences*. 3:1041.
Shin, S.Y., Kim, M.H., Kim, Y.H., Park, H.M. and Yoon, H.S. 2013. Co-expression of monodehydroascorbate reductase and dehydroascorbate reductase from *Brassica rapa* effectively confers tolerance to freezing-induced oxidative stress. *Molecules and Cells*. 36(4):304-315.
Siddique, M.R.B., Hamid, A. and Islam, M.S. 2000. Drought stress effects on water relations of wheat. *Botanical Bulletin of Academia Sinica*. 41:35-39.
Smirnoff, N. 1993. The role of active oxygen in the response of plants to water deficit and desiccation. *New Phytologist*. 125(1):27-58.
Smith, A.D. 2000. Oxford Dictionary of Biochemistry and Molecular Biology. Oxford University Press.
Spaenij-Dekking, L., Kooy-Winkelaar, Y. and Koning, F. 2005. The Ethiopian cereal tef in celiac disease. *New England Journal of Medicine*. 353(16):1748-1749.
Spurr, A.R. 1969. A low-viscosity epoxy resin embedding medium for electron microscopy. *Journal of Ultrastructural Research*. 26(1):31-43.
Sreenivasulu, N., Ramanjulu, S., Ramachandra-Kini, K., Prakash, H.S., Shekar-Shetty, H., Savithri, H.S. and Sudhakar, C. 1999. Total peroxidase activity and peroxidase isoforms as modified by salt stress in two cultivars of fox-tail millet with differential salt tolerance. *Plant Science*. 141(1):1-9.
Staiger, D. 2015. Shaping the *Arabidopsis* transcriptome through alternative splicing. *Advances in Botany*. 2015:13.
Stajich, J.E., Block, D., Boulez, K., Brenner, S.E., Chervitz, S.A., Dagdigian, C., Fuellen, G., Gilbert, J.G., Korf, I., Lapp, H., Lehvaslaiho, H., Matsalla, C., Mungall, C.J., Osborne, B.I., Pocock, M.R., Schattner, P., Senger, M., Stein, L.D., Stupka, E., Wilkinson, M.D. and Birney, E. 2002. The Bioperl toolkit: Perl modules for the life sciences. *Genome Research*. 12(10):1611-1618.
Stallkneecht, G.F., Gilberson, K.M. and Eckhoff, J.L. 1993. Tef: Food crop for humans and animals. *In* New Crops. J. Janick and J.E. Simon, editors. Wiley, New York. 231-234.
Sullivan, D.T., MacIntyre, R., Fuda, N., Fiori, J., Barrilla, J. and Ramizel, L. 2003. Analysis of glycolytic enzyme co-localization in *Drosophila* flight muscle. *Journal of Experimental Biology* 206(Pt 12):2031-2038.
Szarka, A., Tomasskovics, B. and Bánhegyi, G. 2012. The ascorbate-glutathione-α-tocopherol triad in abiotic stress response. *International Journal of Molecular Sciences*. 13(4):4458-4483.
Tadele, Z. and Assefa, K. 2012. Increasing food production in Africa by boosting the productivity of understudied crops. *Agronomy*. 2(4):240-283.
Tadele, Z., Esfeld, K. and Plaza, S. 2010. Applications of high-throughput techniques to the understudied crops of Africa. *Aspects of Applied Biology*. 96:233-240.
Tadesse, D. 1993. Study on genetic variation of landraces of tef [*Eragrostis tef* (Zucc.) Trotter] in Ethiopia. *Genetic Resources and Crop Evolution*. 40(2):101-104.
Takele, A. 1997. Genotypic variability in dry matter production, partitioning and grain yield of tef [*Eragrostis tef* (Zucc.) Trotter] under moisture deficit. *SINET: Ethiopian Journal of Science*. 20(2):177-188.

Takele, A. and Farrant, J.M. 2013. Water relations, gas exchange characteristics and water use efficiency in maize and sorghum after exposure to and recovery from pre and post-flowering dehydration. *African Journal of Agricultural Research* 8(49):6468-6478.
Takele, A., Kebede, H. and Simane, B. 2001. Physiological research in tef. *In* Narrowing The Rift, Tef Research and Development, Proceedings of the "International workshop on Tef Genetics and Improvement", Debre Zeit, Ethiopia. 16-19 October 2000 H. Tefara, G. Belay and M.E. Sorrells, editors, Ethiopia. 177-189.
Tanimoto, T. and Itoh, R. 2001. Changes in leaf rolling rate of the rice and effect of evaporative condition (Crop Physiology and Cell Biology). *Japanese journal of crop science*. 70(3):425-431.
Tatham, A.S., Fido, R.J., Moore, C.M., Kasarda, D.D., Kuzmicky, D.D., Keen, J.N. and Shewry, P.R. 1996. Characterisation of the major prolamins of tef (*Eragrostis tef*) and finger millet (*Eleusine coracana*). *Journal of Cereal Science*. 24:65-71.
Tefara, H. and Belay, G. 2006. *Eragrostis tef* (Zuccagni) Trotter. *In* Plant Resources of Tropical Africa 1: Cereals and pulses. M. Brink and G. Belay, editors. PROTA Foundation, Wageningen, Netherlands. 68-72.
Tefera, H., Assefa, K., Hundera, F., Kefyalew, T. and Teferra, T. 2003. Heritability and genetic advance in recombinant inbred lines of tef (*Eragrostis tef*). *Euphytica*. 131(1):91-96.
Teixeira, J. and Pereira, S. 2007. High salinity and drought act on an organ-dependent manner on potato glutamine synthetase expression and accumulation. *Environmental and Experimental Botany*. 60(1):121-126.
Tesema, A. 2013. Genetic diversity of tef in Ethiopia. *In* Achievements and Prospects of Tef Improvement. A. Assefa, S. Chanyalew and A. Tadele, editors. EIAR-University of Bern, Bern, Switzerland. 15-20.
The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 408(6814):796-815.
Thelen, J.J. 2007. Introduction to proteomics: A brief historical perspective on contemporary approaches. Springer-Verlag, Berlin.
Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., Selbig, J., Muller, L.A., Rhee, S.Y. and Stitt, M. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *The Plant Journal*. 37(6):914-939.
Thompson, A.J., Abu, M. and Hanger, D.P. 2012. Key issues in the acquisition and analysis of qualitative and quantitative mass spectrometry data for peptide-centric proteomic experiments. *Amino Acids*. 43(3):1075-1085.
Turner, N.C. 1979. Drought resistance and adaptation to water deficits in crop plants. *In* Stress Physiology in Crop Plants H. Mussel and R.C. Staples, editors. Wiley, New York. 344-372.
Uematsu, K., Suzuki, N., Iwamae, T., Inui, M. and Yukawa, H. 2012. Increased fructose 1,6-bisphosphate aldolase in plastids enhances growth and photosynthesis of tobacco plants. *Journal of Experimental Botany*. 63(8):3001-3009.
Unlu, M., Morgan, M.E. and Minden, J.S. 1997. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis*. 18(11):2071-2077.
Usadel, B., Nagel, A., Thimm, O., Redestig, H., Blaesing, O.E., Palacios-Rojas, N., Selbig, J., Hannemann, J., Piques, M.C., Steinhauser, D., Scheible, W.-R., Gibon, Y., Morcuende, R., Weicht, D., Meyer, S. and Stitt, M. 2005. Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiology*. 138(3):1195-1204.
Usadel, B., Poree, F., Nagel, A., Lohse, M., Czedik-Eysenberg, A. and Stitt, M. 2009. A guide to using MapMan to visualize and compare omics data in plants: a case study in the crop species, maize. *Plant, Cell and Environment*. 32(9):1211-1229.
Valyova, M., Stoyanov, S., Markovska, Y. and Ganeva, Y. 2012. Evaluation of *in vitro* antioxidant activity and free radical scavenging potential of variety of *Tagetes erecta* L. flowers growing in Bulgaria. *International Journal of Applied Research in Natural Products*. 5(2):19-25.
Vanderschuren, H., Lentz, E., Zainuddin, I. and Gruissem, W. 2013. Proteomics of model and crop plant species: Status, current limitations and strategic advances for crop improvement. *Journal of Proteomics*. 93(0):5-19.
Vavilov, I. 1951. The origin, variation, immunity and breeding of cultivated plants. Translated from Russian by K. Sarrchester, Ronald Press Co, New York, USA.
Vernoud, V., Horton, A.C., Yang, Z. and Nielsen, E. 2003. Analysis of the small GTPase gene superfamily of *Arabidopsis*. *Plant Physiology*. 131(3):1191-1208.
Vincent, D., Ergül, A., Bohlman, M.C., Tattersall, E.A.R., Tillett, R.L., Wheatley, M.D., Woolsey, R., Quilici, D.R., Joets, J., Schlauch, K., Schooley, D.A., Cushman, J.C. and Cramer, G.R. 2007. Proteomic analysis reveals differences between *Vitis vinifera* L. cv. Chardonnay and cv. Cabernet Sauvignon and their responses to water deficit and salinity. *Journal of Experimental Botany*. 58(7):1873-1892.
Wang, Y.Q., Yang, Y., Fei, Z., Yuan, H., Fish, T., Thannhauser, T.W., Mazourek, M., Kochian, L.V., Wang, X. and Li, L. 2013. Proteomic analysis of chromoplasts from six crop species reveals insights into chromoplast function and development. *Journal of Experimental Botany*. 64(4):949-961.

Wang, Z.Q., Xu, X.Y., Gong, Q.Q., Xie, C., Fan, W., Yang, J.L., Lin, Q.S. and Zheng, S.J. 2014. Root proteome of rice studied by iTRAQ provides integrated insight into aluminum stress tolerance mechanisms in plants. *Journal of Proteomics*. 98:189-205.
Waraich, E.A., Ahmad, R., Halim, A. and Aziz, T. 2012. Alleviation of temperature stress by nutrient management in crop plants: a review. *Journal of soil science and plant nutrition*. 12:221-244.
Wasinger, V.C., Zeng, M. and Yau, Y. 2013. Current status and advances in quantitative proteomic mass spectrometry. *Interanational Journal of Proteomics*. 2013:180605.
Whittaker, A., Martinelli, T., Farrant, J.M., Bochicchio, A. and Vazzana, C. 2007. Sucrose phosphate synthase activity and the co-ordination of carbon partitioning during sucrose and amino acid accumulation in desiccation-tolerant leaf material of the C4 resurrection plant *Sporobolus stapfianus* during dehydration. *Journal of Experimental Botany*. 58(13):3775-3787.
Wiese, S., Reidegeld, K.A., Meyer, H.E. and Warscheid, B. 2007. Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics*. 7(3):340-350.
Wilson, P.B., Estavillo, G.M., Field, K.J., Pornsiriwong, W., Carroll, A.J., Howell, K.A., Woo, N.S., Lake, J.A., Smith, S.M., Harvey Millar, A., von Caemmerer, S. and Pogson, B.J. 2009. The nucleotidase/phosphatase *SAL1* is a negative regulator of drought tolerance in *Arabidopsis*. *Plant Journal*. 58(2):299-317.
Wisniewski, J.R., Zougman, A., Nagaraj, N. and Mann, M. 2009. Universal sample preparation method for proteome analysis. *Nature Methods*. 6(5):359-362.
Woyessa, D. and Assefa, F. 2011. Effect of plant growth promoting rhizobacteria on growth and yield of tef (*Eragrostis tef* Zucc. Trotter) under greenhouse condition. *Research Journal of Microbiology*. 6(4):343-355.
Wu, Y. and Cosgrove, D.J. 2000. Adaptation of roots to low water potentials by changes in cell wall extensibility and cell wall proteins. *Journal of Experimental Botany*. 51(350):1543-1553.
Xiao, B., Huang, Y., Tang, N. and Xiong, L. 2007. Over-expression of a LEA gene in rice improves drought resistance under the field conditions. *Theoretical and Applied Genetics*. 115(1):35-46.
Yang, L., Jiang, T., Fountain, J.C., Scully, B.T., Lee, R.D., Kemerait, R.C., Chen, S. and Guo, B. 2014. Protein profiles reveal diverse responsive signaling pathways in kernels of two maize inbred lines with contrasting drought sensitivity. *International Journal of Molecular Sciences*. 15(10):18892-18918.
Yu, J.K., Graznak, E., Breseghello, F., Tefera, H. and Sorrells, M.E. 2007. QTL mapping of agronomic traits in tef [*Eragrostis tef* (Zucc) Trotter]. *BMC Plant Biology*. 7:30.
Yu, J.K., Kantety, R.V., Graznak, E., Benscher, D., Tefera, H. and Sorrells, M.E. 2006. A genetic linkage map for tef [*Eragrostis tef* (Zucc.) Trotter]. *Theoretical and Applied Genetics*. 113(6):1093-1102.
Zarsky, V., Cvrckova, F., Bischoff, F. and Palme, K. 1997. At-GDI1 from *Arabidopsis thaliana* encodes a rab-specific GDP dissociation inhibitor that complements the sec19 mutation of *Saccharomyces cerevisiae*. *FEBS Letters*. 403(3):303-308.
Zeid, M., Assefa, K., Haddis, A., Chanyalew, S. and Sorrells, M.E. 2012. Genetic diversity in tef (*Eragrostis tef*) germplasm using SSR markers. *Field Crops Research*. 127(0):64-70.
Zeid, M., Belay, G., Mulkey, S., Poland, J. and Sorrells, M.E. 2011. QTL mapping for yield and lodging resistance in an enhanced SSR-based map for tef. *Theoretical and Applied Genetics*. 122(1):77-93.
Zewdu, A.D. and Solomon, W.K. 2007. Moisture-dependent physical properties of tef seed. *Biosystems Engineering*. 96(1):57-63.
Zhang, D., Ayele, M., Tefera, H. and Nguyen, H.T. 2001. RFLP linkage map of the Ethiopian cereal tef [*Eragrostis tef* (Zucc) Trotter]. *Theoretical and Applied Genetics*. 102(6-7):957-964.
Zhang, J. and Kirkham, M.B. 1994. Drought-stress-induced changes in activities of superoxide dismutase, catalase, and peroxidase in wheat species. *Plant and Cell Physiology*. 35(5):785-791.
Zhang, W., Gruszewski, H.A., Chevone, B.I. and Nessler, C.L. 2008. An Arabidopsis purple acid phosphatase with phytase activity increases foliar ascorbate. *Plant Physiology*. 146(2):431-440.
Zhu, L.-Q., Zheng, H.Y., Peng, C.X., Liu, D., Li, H.L., Wang, Q. and Wang, J.Z. 2010. Protein phosphatase 2A facilitates axonogenesis by dephosphorylating CRMP2. *The Journal of Neuroscience*. 30(10):3839-3848.
Zhu, Q., Smith, S.M., Ayele, M., Yang, L., Jogi, A., Chaluvadi, S.R. and Bennetzen, J.L. 2012. High-throughput discovery of mutations in tef semi-dwarfing genes by next-generation sequencing analysis. *Genetics*. 192(3):819-829.